

Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/US2005/010912

International filing date: 31 March 2005 (31.03.2005)

Document type: Certified copy of priority document

Document details: Country/Office: US
Number: 60/559,042
Filing date: 01 April 2004 (01.04.2004)

Date of receipt at the International Bureau: 12 August 2005 (12.08.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)



World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse

1352391

UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

August 02, 2005

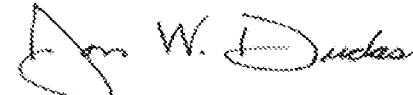
THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM
THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK
OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT
APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A
FILING DATE.

APPLICATION NUMBER: 60/559,042

FILING DATE: April 01, 2004

RELATED PCT APPLICATION NUMBER: PCT/US05/10912

Certified by



Under Secretary of Commerce
for Intellectual Property
and Director of the United States
Patent and Trademark Office



01919 U.S.PTO
040104

PTO/SB/16 (08-03)

Approved for use through 7/31/2006. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

PROVISIONAL APPLICATION FOR PATENT COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53(c).

Express Mail Label No. EL 961008104 US

22857 U.S.PTO
60/559042

040104

INVENTOR(S)		
Given Name (first and middle [if any]) Steven Andreas Stefan M.	Family Name or Surname MAH BRAUN KAMMERER	Residence (City and either State or Foreign Country) San Diego, California San Diego, California San Diego, California
Additional inventors are being named on the <u>1</u> separately numbered sheets attached hereto		
TITLE OF THE INVENTION (500 characters max) METHODS FOR IDENTIFYING RISK OF OSTEOARTHRITIS AND TREATMENTS THEREOF		
Direct all correspondence to: CORRESPONDENCE ADDRESS		
<input checked="" type="checkbox"/> Customer Number:	25225	
OR		
<input type="checkbox"/> Firm or Individual Name		
Address		
City	State	Zip
Country	Telephone	Fax
ENCLOSED APPLICATION PARTS (check all that apply)		
<input checked="" type="checkbox"/> Specification Number of Pages	162	<input type="checkbox"/> CD(s), Number
<input checked="" type="checkbox"/> Drawing(s) Number of Sheets	1	<input checked="" type="checkbox"/> Other
<input checked="" type="checkbox"/> Application Data Sheet. See 37 CFR 1.76 (4 pages) (specify): Return Receipt Postcard		
METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT		
<input checked="" type="checkbox"/> Applicant claims small entity status. See 37 CFR 1.27.	FILING FEE AMOUNT (\$)	
<input type="checkbox"/> A check or money order is enclosed to cover the filing fees.		
<input checked="" type="checkbox"/> The Director is hereby authorized to charge filing fees or credit any overpayment to Deposit Account Number: <u>03-1952</u>	80.00	
<input type="checkbox"/> Payment by credit card. Form PTO-2038 is attached.		
The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.		
<input checked="" type="checkbox"/> No	<input type="checkbox"/> Yes, the name of the U.S. Government agency and the Government contract number are: _____	

Respectfully submitted,

[Page 1 of 2]

Date April 1, 2004

SIGNATURE

TYPED OR
PRINTED NAME

TELEPHONE

Bruce D. Grant

(858) 720-7962

REGISTRATION NO.
(if appropriate)

47,608

Docket Number: 524593008900**USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT**

I hereby certify that this correspondence is being deposited with the U.S. Postal Service as Express Mail, Airbill No. EL 961008104 US, in an envelope addressed to: Mail Stop Provisional Patent Application, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450, on the date shown below.

Dated: 4/1/04Signature: Deborah Wykes (Deborah Wykes)

PROVISIONAL APPLICATION COVER SHEET
Additional Page

PTO/SB/16 (08-03)

Approved for use through 07/31/06. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

Docket Number 524593008900

INVENTOR(S)/APPLICANT(S)

Given Name (first and middle if any)	Family or Surname	Residence (City and either State or Foreign Country)
Matthew Roberts	NELSON	San Marcos, California
Rikard Henry	RENELAND	San Diego, California
Maria L.	LANGDOWN	San Diego, California

[Page 2 of 2]

**METHODS FOR IDENTIFYING RISK OF OSTEOARTHRITIS AND
TREATMENTS THEREOF**

Field of the Invention

[0001] The invention relates to genetic methods for identifying risk of osteoarthritis and treatments that specifically target such diseases.

Background

[0002] Osteoarthritis (OA) is a chronic disease usually affecting weight-bearing synovial joints. There are approximately 20 million Americans affected by OA and it is the leading cause of disability in the United States. In addition to extensive human suffering, OA also accounts for nearly all knee replacements and more than half of all hip replacements in the United States. Despite its prevalence, OA is poorly understood and there are few treatments available besides anti-inflammatory drugs and joint replacement.

[0003] Most commonly affecting middle-aged and older people, OA can range from very mild to very severe. It affects hands and weight-bearing joints such as knees, hips, feet and the back. Knee OA can be as disabling as any cardiovascular disease except stroke.

[0004] OA is characterized by the breakdown of cartilage in joints. Cartilage in joints cushions the ends of bones, and cartilage breakdown causes bones to rub against each other, causing pain and loss of movement. Type II collagen is the main component of cartilage, comprising 15-25% of the wet weight, approximately half the dry weight, and representing 90-95% of the total collagen content in the tissue. It forms fibrils that endow cartilage with tensile strength (Mayne, R. Arthritis Rhuem. 32:241-246 (1989)).

Summary

[0005] It has been discovered that certain polymorphic variations in human genomic DNA are associated with osteoarthritis. In particular, polymorphic variants in a locus containing a *ADAMTS2* region in human genomic DNA have been associated with risk of osteoarthritis.

[0006] Thus, featured herein are methods for identifying a subject at risk of osteoarthritis and/or a risk of osteoarthritis in a subject, which comprise detecting the presence or absence of one or more polymorphic variations associated with osteoarthritis in or around the loci described herein in a human nucleic acid sample. In an embodiment, two or more polymorphic variations are detected in two or more regions of which one is the *ADAMTS2* region. In certain embodiments, 3 or more, or 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 or more polymorphic variants are detected.

[0007] Also featured are nucleic acids that include one or more polymorphic variations associated with occurrence of osteoarthritis, as well as polypeptides encoded by these nucleic acids. In addition, provided are methods for identifying candidate therapeutic molecules for treating osteoarthritis, as well as methods for treating osteoarthritis in a subject by identifying a subject at risk of osteoarthritis and treating the subject with a suitable prophylactic, treatment or therapeutic molecule.

[0008] Also provided are compositions comprising a cell from a subject having osteoarthritis or at risk of osteoarthritis and/or a *ADAMTS2* nucleic acid, with a RNAi, siRNA, antisense DNA or RNA, or ribozyme nucleic acid designed from a *ADAMTS2* nucleotide sequence. In an embodiment, the RNAi, siRNA, antisense DNA or RNA, or ribozyme nucleic acid is designed from a *ADAMTS2* nucleotide sequence that includes one or more polymorphic variations associated with osteoarthritis, and in some instances, specifically interacts with such a nucleotide sequence. Further, provided are arrays of nucleic acids bound to a solid surface, in which one or more nucleic acid molecules of the array have a *ADAMTS2* nucleotide sequence, or a fragment or substantially identical nucleic acid thereof, or a complementary nucleic acid of the foregoing. Featured also are compositions comprising a cell from a subject having osteoarthritis or at risk of osteoarthritis and/or a *ADAMTS2* polypeptide, with an antibody that specifically binds to the polypeptide. Thus, featured is an antibody that specifically binds to an epitope in the polypeptide that includes an amino acid encoded by a polymorphic site associated with osteoarthritis. In certain embodiments, the antibody specifically binds to an epitope comprising a valine or isoleucine encoded by rs398829 (e.g., an antibody that binds to an epitope comprising a valine at position 245 in an *ADAMTS2* polypeptide) A valine at position 245 is associated with increased risk of osteoarthritis.

Brief Description of the Drawings

[0009] Figure 1 shows proximal SNPs in a *ADAMTS2* region in genomic DNA. The position of each SNP in the chromosome is shown on the x-axis and the y-axis provides the negative logarithm of the p-value comparing the estimated allele to that of the control group. Also shown in the figure are exons and introns of the region in the approximate chromosomal positions.

Detailed Description

[0010] It has been discovered that a polymorphic variant in a locus containing a *ADAMTS2* region is associated with occurrence of osteoarthritis in subjects. Thus, detecting genetic determinants associated with an increased risk of osteoarthritis occurrence can lead to early identification of a predisposition to osteoarthritis and early prescription of preventative measures. Also, associating a *ADAMTS2* polymorphic variant with osteoarthritis has provided new targets for screening molecules useful in treatments of osteoarthritis.

Osteoarthritis and Sample Selection

[0011] Osteoarthritis (OA), or degenerative joint disease, is one of the oldest and most common types of arthritis. It is characterized by the breakdown of the joint's cartilage. Cartilage is the part of the joint that cushions the ends of bones, and its breakdown causes bones to rub against each other, causing pain and loss of movement. Type II collagen is the main component of cartilage, comprising 15-25% of the wet weight, approximately half the dry weight, and representing 90-95% of the total collagen content in the tissue. It forms fibrils that endow cartilage with tensile strength (Mayne, R. Arthritis Rheum. 32:241-246 (1989)).

[0012] Most commonly affecting middle-aged and older people, OA can range from very mild to very severe. It affects hands and weight-bearing joints such as knees, hips, feet and the back. Knee OA can be as disabling as any cardiovascular disease except stroke. Whereas Ehlers-Danlos syndrome type VIIC is characterized by the retention of the N-terminal propeptide of type I collagen, osteoarthritis has been characterized by increased levels of type II collagen in osteoarthritic cartilage as measured by elevated C-propeptide concentrations (Nelson et al. (1998) J. Clin. Invest. 102(12):2115-2125).

[0013] Osteoarthritis affects an estimated 20.7 million Americans, mostly after age 45, with women more commonly affected than men. Physicians make a diagnosis of OA based on a physical exam and history of symptoms. X-rays are used to confirm diagnosis. Most people over 60 reflect the disease on X-ray, and about one-third have actual symptoms.

[0014] There are many factors that can cause OA. Obesity may lead to osteoarthritis of the knees. In addition, people with joint injuries due to sports, work-related activity or accidents may be at increased risk of developing OA.

[0015] Genetics has a role in the development of OA. Some people may be born with defective cartilage or with slight defects in the way that joints fit together. As a person ages, these defects may cause early cartilage breakdown in the joint or the inability to repair damaged or deteriorated cartilage in the joint.

[0016] Inclusion or exclusion of samples for an osteoarthritis pool may be based upon the following criteria: ethnicity (e.g., samples derived from an individual characterized as Caucasian); parental ethnicity (e.g., samples derived from an individual of British paternal and maternal descent); relevant phenotype information for the individual (e.g., case samples derived from individuals diagnosed with specific knee osteoarthritis (OA) and were recruited from an OA knee replacement clinic). Control samples may be selected based on relevant phenotype information for the individual (e.g., derived from individuals free of OA at several sites (knee, hand, hip etc)); and no family history of OA and/or rheumatoid arthritis. Additional phenotype information collected for both cases and controls may include age of the individual, gender, family history of OA, diagnosis with osteoarthritis (joint location of OA, date of primary diagnosis, age of individual as of primary diagnosis), knee history (current symptoms,

any major knee injury, meniscectomy, knee replacement surgery, age of surgery), HRT history, osteoporosis diagnosis.

[0017] Based in part upon selection criteria set forth above, individuals having osteoarthritis can be selected for genetic studies. Also, individuals having no history of osteoarthritis often are selected for genetic studies, as described hereafter.

Polymorphic Variants Associated with Osteoarthritis

[0018] A genetic analysis provided herein linked osteoarthritis with polymorphic variant nucleic acid sequences in the human genome. As used herein, the term “polymorphic site” refers to a region in a nucleic acid at which two or more alternative nucleotide sequences are observed in a significant number of nucleic acid samples from a population of individuals. A polymorphic site may be a nucleotide sequence of two or more nucleotides, an inserted nucleotide or nucleotide sequence, a deleted nucleotide or nucleotide sequence, or a microsatellite, for example. A polymorphic site that is two or more nucleotides in length may be 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 or more, 20 or more, 30 or more, 50 or more, 75 or more, 100 or more, 500 or more, or about 1000 nucleotides in length, where all or some of the nucleotide sequences differ within the region. A polymorphic site is often one nucleotide in length, which is referred to herein as a “single nucleotide polymorphism” or a “SNP.”

[0019] Where there are two, three, or four alternative nucleotide sequences at a polymorphic site, each nucleotide sequence is referred to as a “polymorphic variant” or “nucleic acid variant.” Where two polymorphic variants exist, for example, the polymorphic variant represented in a minority of samples from a population is sometimes referred to as a “minor allele” and the polymorphic variant that is more prevalently represented is sometimes referred to as a “major allele.” Many organisms possess a copy of each chromosome (e.g., humans), and those individuals who possess two major alleles or two minor alleles are often referred to as being “homozygous” with respect to the polymorphism, and those individuals who possess one major allele and one minor allele are normally referred to as being “heterozygous” with respect to the polymorphism. Individuals who are homozygous with respect to one allele are sometimes predisposed to a different phenotype as compared to individuals who are heterozygous or homozygous with respect to another allele.

[0020] In genetic analysis that associate polymorphic variants with osteoarthritis, samples from individuals having osteoarthritis and individuals not having osteoarthritis often are allelotyped and/or genotyped. The term “allelotype” as used herein refers to a process for determining the allele frequency for a polymorphic variant in pooled DNA samples from cases and controls. By pooling DNA from each group, an allele frequency for each SNP in each group is calculated. These allele frequencies are then compared to one another. The term “genotyped” as used herein refers to a process for determining a

genotype of one or more individuals, where a “genotype” is a representation of one or more polymorphic variants in a population.

[0021] A genotype or polymorphic variant may be expressed in terms of a “haplotype,” which as used herein refers to two or more polymorphic variants occurring within genomic DNA in a group of individuals within a population. For example, two SNPs may exist within a gene where each SNP position includes a cytosine variation and an adenine variation. Certain individuals in a population may carry one allele (heterozygous) or two alleles (homozygous) having the gene with a cytosine at each SNP position. As the two cytosines corresponding to each SNP in the gene travel together on one or both alleles in these individuals, the individuals can be characterized as having a cytosine/cytosine haplotype with respect to the two SNPs in the gene.

[0022] As used herein, the term “phenotype” refers to a trait which can be compared between individuals, such as presence or absence of a condition, a visually observable difference in appearance between individuals, metabolic variations, physiological variations, variations in the function of biological molecules, and the like. An example of a phenotype is occurrence of osteoarthritis.

[0023] Researchers sometimes report a polymorphic variant in a database without determining whether the variant is represented in a significant fraction of a population. Because a subset of these reported polymorphic variants are not represented in a statistically significant portion of the population, some of them are sequencing errors and/or not biologically relevant. Thus, it is often not known whether a reported polymorphic variant is statistically significant or biologically relevant until the presence of the variant is detected in a population of individuals and the frequency of the variant is determined. Methods for detecting a polymorphic variant in a population are described herein, specifically in Example 2. A polymorphic variant is statistically significant and often biologically relevant if it is represented in 5% or more of a population, sometimes 10% or more, 15% or more, or 20% or more of a population, and often 25% or more, 30% or more, 35% or more, 40% or more, 45% or more, or 50% or more of a population.

[0024] A polymorphic variant may be detected on either or both strands of a double-stranded nucleic acid. Also, a polymorphic variant may be located within an intron or exon of a gene or within a portion of a regulatory region such as a promoter, a 5' untranslated region (UTR), a 3' UTR, and in DNA (e.g., genomic DNA (gDNA) and complementary DNA (cDNA)), RNA (e.g., mRNA, tRNA, and rRNA), or a polypeptide. Polymorphic variations may or may not result in detectable differences in gene expression, polypeptide structure, or polypeptide function.

[0025] It was determined that polymorphic variations associated with an increased risk of osteoarthritis existed in a *ADAMTS2* region in SEQ ID NO: 1. In certain embodiments, a polymorphic variant at position rs398829 in the human genome was associated with an increased risk of osteoarthritis, and in a specific embodiment, a guanine at position rs398829 was associated with an increased risk of osteoarthritis.

[0026] Polymorphic variants in and around the *ADAMTS2* locus were tested for association with osteoarthritis. These include polymorphic variants at positions in SEQ ID NO: 1 selected from the group consisting of 210, 3608, 3609, 4318, 5593, 5629, 5639, 5640, 8943, 17968, 19887, 21034, 21085, , 21596, 23379, 23432, 24007, 26121, 26273, 26755, 27411, 27710, 27842, 28379, 29603, 31232, 31504, 32583, 32794, 32840, 33044, 33150, 33218, 33513, 33959, 34486, 36289, 36570, 38247, 38477, 38518, 38529, 38667, 39781, 39856, 39927, 40506, 41869, 42452, 44788, 46059, 46846, 47712, 48796, 49441, 49602, 49723, 50050, 50171, 50477, 50818, 50833, 50881, 50882, 51386, 51534, 52317, 52368, 52970, 53023, 53356, 53882, 54553, 55475, 55530, 55691, 55848, 55879, 56316, 56911, 57320, 57391, 57437, 57478, 57500, 59111, 59333, 59715, 59804, 59851, 59929, 60052, 60240, 60359, 60381, 60456, 60724, 60875, 60968, 60978, 60998, 61557, 62091, 62645, 62943, 63131, 63145, 63406, 63427, 63554, 63661, 64093, 64153, 64409, 64544, 65257, 65626, 65739 , 66392, 66720, 69177, 69336, 69636, 69823, 69928, 70547, 70633, 71805, 72181, 72200, 72474, 72567, 72973, 73468, 73889, 75730, 75970, 76114, 76342, 76449, 76465, 76791, 78042, 80758, 80778, 81356, 81576, 81689, 81759, 81950, 82562, 83591, 83700, 83821, 83842, 83923, 83929, 84021, 84175, 84417, 84747, 85746, 86129, 86335, 87315, 87648, 87764, 87770, 88221, 90474, 91148, 91150, 91160, 91733, 91772, 91785, 93140, 93148, 96080, 96157, 96313, 96759, 97026, 97320, 97732, 98713, 99707, 99959, 100009, 100020, 100065, 100086, 101270, 101276, 101371, 101376, 101439, 101820, 102392, 102602, 102604, 102896, 189104, 189134 and 189205.

Polymorphic variants at the following positions in SEQ ID NO: 1 in particular were associated with an increased risk of osteoarthritis: 5640, 33150, 38247, 38529, 46846, 49723, 50050, 63427, 73889, 189104 and rs428901, where specific embodiments are directed to positions 46846, 73889, 189104 and/or rs428901. In particular, the following polymorphic variants in SEQ ID NO: 1 were associated with risk of osteoarthritis: a cytosine at position 5640, a cytosine at position 33150, an adenine at position 38247, a thymine at position 38529, an adenine at position 46846, a cytosine at position 49723, a cytosine at position 50050, a cytosine a position 63427, a guanine at position 73889, a thymine at position 189104, and an adenine at position rs428901.

[0027] Based in part upon analyses summarized in Figure 1, a region with significant association has been identified in a locus associated with osteoarthritis. Any polymorphic variants associated with osteoarthritis in a region of significant association can be utilized for embodiments described herein. For example, polymorphic variants in a region spanning chromosome positions 178746000 to 178751000 (approximately 5,000 nucleotides in length) in a *ADAMTS2* locus have significant association (chromosome positions are within NCBI's Genome build 34).

Additional Polymorphic Variants Associated with Osteoarthritis

[0028] Also provided is a method for identifying polymorphic variants proximal to an incident, founder polymorphic variant associated with osteoarthritis. Thus, featured herein are methods for

identifying a polymorphic variation associated with osteoarthritis that is proximal to an incident polymorphic variation associated with osteoarthritis, which comprises identifying a polymorphic variant proximal to the incident polymorphic variant associated with osteoarthritis, where the incident polymorphic variant is in a *ADAMTS2* nucleotide sequence. The nucleotide sequence often comprises a polynucleotide sequence selected from the group consisting of (a) a polynucleotide sequence of SEQ ID NO: 1-3; (b) a polynucleotide sequence that encodes a polypeptide having an amino acid sequence encoded by a polynucleotide sequence of SEQ ID NO: 1-3; and (c) a polynucleotide sequence that encodes a polypeptide having an amino acid sequence that is 90% or more identical to an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1-3 or a polynucleotide sequence 90% or more identical to the polynucleotide sequence of SEQ ID NO: 1-3. The presence or absence of an association of the proximal polymorphic variant with osteoarthritis then is determined using a known association method, such as a method described in the Examples hereafter. In an embodiment, the incident polymorphic variant is a polymorphic variant associated with osteoarthritis described herein. In another embodiment, the proximal polymorphic variant identified sometimes is a publicly disclosed polymorphic variant, which for example, sometimes is published in a publicly available database. In other embodiments, the polymorphic variant identified is not publicly disclosed and is discovered using a known method, including, but not limited to, sequencing a region surrounding the incident polymorphic variant in a group of nucleic samples. Thus, multiple polymorphic variants proximal to an incident polymorphic variant are associated with osteoarthritis using this method.

[0029] The proximal polymorphic variant often is identified in a region surrounding the incident polymorphic variant. In certain embodiments, this surrounding region is about 50 kb flanking the first polymorphic variant (e.g. about 50 kb 5' of the first polymorphic variant and about 50 kb 3' of the first polymorphic variant), and the region sometimes is composed of shorter flanking sequences, such as flanking sequences of about 40 kb, about 30 kb, about 25 kb, about 20 kb, about 15 kb, about 10 kb, about 7 kb, about 5 kb, or about 2 kb 5' and 3' of the incident polymorphic variant. In other embodiments, the region is composed of longer flanking sequences, such as flanking sequences of about 55 kb, about 60 kb, about 65 kb, about 70 kb, about 75 kb, about 80 kb, about 85 kb, about 90 kb, about 95 kb, or about 100 kb 5' and 3' of the incident polymorphic variant.

[0030] In certain embodiments, polymorphic variants associated with osteoarthritis are identified iteratively. For example, a first proximal polymorphic variant is associated with osteoarthritis using the methods described above and then another polymorphic variant proximal to the first proximal polymorphic variant is identified (e.g., publicly disclosed or discovered) and the presence or absence of an association of one or more other polymorphic variants proximal to the first proximal polymorphic variant with osteoarthritis is determined.

[0031] The methods described herein are useful for identifying or discovering additional polymorphic variants that may be used to further characterize a gene, region or loci associated with a condition, a disease (*e.g.*, osteoarthritis), or a disorder. For example, allelotyping or genotyping data from the additional polymorphic variants may be used to identify a functional mutation or a region of linkage disequilibrium. In certain embodiments, polymorphic variants identified or discovered within a region comprising the first polymorphic variant associated with osteoarthritis are genotyped using the genetic methods and sample selection techniques described herein, and it can be determined whether those polymorphic variants are in linkage disequilibrium with the first polymorphic variant. The size of the region in linkage disequilibrium with the first polymorphic variant also can be assessed using these genotyping methods. Thus, provided herein are methods for determining whether a polymorphic variant is in linkage disequilibrium with a first polymorphic variant associated with osteoarthritis, and such information can be used in prognosis/diagnosis methods described herein.

Isolated Nucleic Acids

[0032] Featured herein are isolated *ADAMTS2* nucleic acid variants depicted in SEQ ID NO: 1-3, and substantially identical nucleic acids thereof. A nucleic acid variant may be represented on one or both strands in a double-stranded nucleic acid or on one chromosomal complement (heterozygous) or both chromosomal complements (homozygous). *ADAMTS2* exists in two forms, a "long" form comprising a molecule approximately 130 kDa in length (*e.g.*, SEQ ID NO: 2 for cDNA sequence and SEQ ID NO: 4 for amino acid sequence), and a "short" form comprising a molecule approximately 70 kDa in length (*e.g.*, SEQ ID NO: 3 for cDNA sequence and SEQ ID NO: 5 for amino acid sequence). Provided herein are polynucleotide sequences encoding both the short and long forms of *ADAMTS2*.

[0033] As used herein, the term "nucleic acid" includes DNA molecules (*e.g.*, a complementary DNA (cDNA) and genomic DNA (gDNA)) and RNA molecules (*e.g.*, mRNA, rRNA, siRNA and tRNA) and analogs of DNA or RNA, for example, by use of nucleotide analogs. The nucleic acid molecule can be single-stranded and it is often double-stranded. The term "isolated or purified nucleic acid" refers to nucleic acids that are separated from other nucleic acids present in the natural source of the nucleic acid. For example, with regard to genomic DNA, the term "isolated" includes nucleic acids which are separated from the chromosome with which the genomic DNA is naturally associated. An "isolated" nucleic acid is often free of sequences which naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and/or 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated nucleic acid molecule can contain less than about 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb or 0.1 kb of 5' and/or 3' nucleotide sequences which flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular

material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized. As used herein, the term “gene” refers to a nucleotide sequence that encodes a polypeptide.

[0034] Also included herein are nucleic acid fragments. These fragments often have a nucleotide sequence identical to a nucleotide sequence of SEQ ID NO: 1-3, a nucleotide sequence substantially identical to a nucleotide sequence of SEQ ID NO: 1-3, or a nucleotide sequence that is complementary to the foregoing. The nucleic acid fragment may be identical, substantially identical or homologous to a nucleotide sequence in an exon or an intron in a nucleotide sequence of SEQ ID NO: 1-3, and may encode a domain or part of a domain of a polypeptide. Sometimes, the fragment will comprises one or more of the polymorphic variations described herein as being associated with osteoarthritis. The nucleic acid fragment is often 50, 100, or 200 or fewer base pairs in length, and is sometimes about 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 2000, 3000, 4000, 5000, 10000, 15000, or 20000 base pairs in length. A nucleic acid fragment that is complementary to a nucleotide sequence identical or substantially identical to a nucleotide sequence in SEQ ID NO: 1-3 and hybridizes to such a nucleotide sequence under stringent conditions is often referred to as a “probe.” Nucleic acid fragments often include one or more polymorphic sites, or sometimes have an end that is adjacent to a polymorphic site as described hereafter.

[0035] An example of a nucleic acid fragment is an oligonucleotide. As used herein, the term “oligonucleotide” refers to a nucleic acid comprising about 8 to about 50 covalently linked nucleotides, often comprising from about 8 to about 35 nucleotides, and more often from about 10 to about 25 nucleotides. The backbone and nucleotides within an oligonucleotide may be the same as those of naturally occurring nucleic acids, or analogs or derivatives of naturally occurring nucleic acids, provided that oligonucleotides having such analogs or derivatives retain the ability to hybridize specifically to a nucleic acid comprising a targeted polymorphism. Oligonucleotides described herein may be used as hybridization probes or as components of prognostic or diagnostic assays, for example, as described herein.

[0036] Oligonucleotides are typically synthesized using standard methods and equipment, such as the ABI™3900 High Throughput DNA Synthesizer and the EXPEDITE™ 8909 Nucleic Acid Synthesizer, both of which are available from Applied Biosystems (Foster City, CA). Analogs and derivatives are exemplified in U.S. Pat. Nos. 4,469,863; 5,536,821; 5,541,306; 5,637,683; 5,637,684; 5,700,922; 5,717,083; 5,719,262; 5,739,308; 5,773,601; 5,886,165; 5,929,226; 5,977,296; 6,140,482; WO 00/56746; WO 01/14398, and related publications. Methods for synthesizing oligonucleotides comprising such analogs or derivatives are disclosed, for example, in the patent publications cited above

and in U.S. Pat. Nos. 5,614,622; 5,739,314; 5,955,599; 5,962,674; 6,117,992; in WO 00/75372; and in related publications.

[0037] Oligonucleotides may also be linked to a second moiety. The second moiety may be an additional nucleotide sequence such as a tail sequence (*e.g.*, a polyadenosine tail), an adapter sequence (*e.g.*, phage M13 universal tail sequence), and others. Alternatively, the second moiety may be a non-nucleotide moiety such as a moiety which facilitates linkage to a solid support or a label to facilitate detection of the oligonucleotide. Such labels include, without limitation, a radioactive label, a fluorescent label, a chemiluminescent label, a paramagnetic label, and the like. The second moiety may be attached to any position of the oligonucleotide, provided the oligonucleotide can hybridize to the nucleic acid comprising the polymorphism.

Uses for Nucleic Acid Sequence

[0038] Nucleic acid coding sequences may be used for diagnostic purposes for detection and control of polypeptide expression. Also, included herein are oligonucleotide sequences such as antisense RNA, small-interfering RNA (siRNA) and DNA molecules and ribozymes that function to inhibit translation of a polypeptide. Antisense techniques and RNA interference techniques are known in the art and are described herein.

[0039] Ribozymes are enzymatic RNA molecules capable of catalyzing the specific cleavage of RNA. The mechanism of ribozyme action involves sequence specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, hammerhead motif ribozyme molecules may be engineered that specifically and efficiently catalyze endonucleolytic cleavage of RNA sequences corresponding to or complementary to *ADAMTS2* nucleotide sequences. Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites which include the following sequences, GUA, GUU and GUC. Once identified, short RNA sequences of between fifteen (15) and twenty (20) ribonucleotides corresponding to the region of the target gene containing the cleavage site may be evaluated for predicted structural features such as secondary structure that may render the oligonucleotide sequence unsuitable. The suitability of candidate targets may also be evaluated by testing their accessibility to hybridization with complementary oligonucleotides, using ribonuclease protection assays.

[0040] Antisense RNA and DNA molecules, siRNA and ribozymes may be prepared by any method known in the art for the synthesis of RNA molecules. These include techniques for chemically synthesizing oligodeoxyribonucleotides well known in the art such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA sequences encoding the antisense RNA molecule. Such DNA sequences may be incorporated

into a wide variety of vectors which incorporate suitable RNA polymerase promoters such as the T7 or SP6 polymerase promoters. Alternatively, antisense cDNA constructs that synthesize antisense RNA constitutively or inducibly, depending on the promoter used, can be introduced stably into cell lines.

[0041] DNA encoding a polypeptide also may have a number of uses for the diagnosis of diseases, including osteoarthritis, resulting from aberrant expression of a target gene described herein. For example, the nucleic acid sequence may be used in hybridization assays of biopsies or autopsies to diagnose abnormalities of expression or function (*e.g.*, Southern or Northern blot analysis, *in situ* hybridization assays).

[0042] In addition, the expression of a polypeptide during embryonic development may also be determined using nucleic acid encoding the polypeptide. As addressed, *infra*, production of functionally impaired polypeptide is the cause of various disease states, such as osteoarthritis. *In situ* hybridizations using polypeptide as a probe may be employed to predict problems related to osteoarthritis. Further, as indicated, *infra*, administration of human active polypeptide, recombinantly produced as described herein, may be used to treat disease states related to functionally impaired polypeptide. Alternatively, gene therapy approaches may be employed to remedy deficiencies of functional polypeptide or to replace or compete with dysfunctional polypeptide.

Expression Vectors, Host Cells, and Genetically Engineered Cells

[0043] Provided herein are nucleic acid vectors, often expression vectors, which contain a *ADAMTS2* nucleotide sequence, or a substantially identical sequence thereof. As used herein, the term “vector” refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked and can include a plasmid, cosmid, or viral vector. The vector can be capable of autonomous replication or it can integrate into a host DNA. Viral vectors may include replication defective retroviruses, adenoviruses and adeno-associated viruses for example.

[0044] A vector can include a *ADAMTS2* nucleotide sequence in a form suitable for expression of an encoded target polypeptide or target nucleic acid in a host cell. A “target polypeptide” is a polypeptide encoded by a *ADAMTS2* nucleotide sequence, or a substantially identical nucleotide sequence thereof. The recombinant expression vector typically includes one or more regulatory sequences operatively linked to the nucleic acid sequence to be expressed. The term “regulatory sequence” includes promoters, enhancers and other expression control elements (*e.g.*, polyadenylation signals). Regulatory sequences include those that direct constitutive expression of a nucleotide sequence, as well as tissue-specific regulatory and/or inducible sequences. The design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of polypeptide desired, and the like. Expression vectors can be introduced into host cells to produce target polypeptides, including fusion polypeptides.

[0045] Recombinant expression vectors can be designed for expression of target polypeptides in prokaryotic or eukaryotic cells. For example, target polypeptides can be expressed in *E. coli*, insect cells (e.g., using baculovirus expression vectors), yeast cells, or mammalian cells. Suitable host cells are discussed further in Goeddel, *Gene Expression Technology: Methods in Enzymology 185*, Academic Press, San Diego, CA (1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

[0046] Expression of polypeptides in prokaryotes is most often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion polypeptides. Fusion vectors add a number of amino acids to a polypeptide encoded therein, usually to the amino terminus of the recombinant polypeptide. Such fusion vectors typically serve three purposes: 1) to increase expression of recombinant polypeptide; 2) to increase the solubility of the recombinant polypeptide; and 3) to aid in the purification of the recombinant polypeptide by acting as a ligand in affinity purification. Often, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant polypeptide to enable separation of the recombinant polypeptide from the fusion moiety subsequent to purification of the fusion polypeptide. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith & Johnson, *Gene 67*: 31-40 (1988)), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding polypeptide, or polypeptide A, respectively, to the target recombinant polypeptide.

[0047] Purified fusion polypeptides can be used in screening assays and to generate antibodies specific for target polypeptides. In a therapeutic embodiment, fusion polypeptide expressed in a retroviral expression vector is used to infect bone marrow cells that are subsequently transplanted into irradiated recipients. The pathology of the subject recipient is then examined after sufficient time has passed (e.g., six (6) weeks).

[0048] Expressing the polypeptide in host bacteria with an impaired capacity to proteolytically cleave the recombinant polypeptide is often used to maximize recombinant polypeptide expression (Gottesman, S., *Gene Expression Technology: Methods in Enzymology, Academic Press, San Diego, California 185*: 119-128 (1990)). Another strategy is to alter the nucleotide sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (Wada *et al.*, *Nucleic Acids Res. 20*: 2111-2118 (1992)). Such alteration of nucleotide sequences can be carried out by standard DNA synthesis techniques.

[0049] When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. Recombinant mammalian expression vectors are

often capable of directing expression of the nucleic acid in a particular cell type (e.g., tissue-specific regulatory elements are used to express the nucleic acid). Non-limiting examples of suitable tissue-specific promoters include an albumin promoter (liver-specific; Pinkert *et al.*, *Genes Dev.* 1: 268-277 (1987)), lymphoid-specific promoters (Calame & Eaton, *Adv. Immunol.* 43: 235-275 (1988)), promoters of T cell receptors (Winoto & Baltimore, *EMBO J.* 8: 729-733 (1989)) promoters of immunoglobulins (Banerji *et al.*, *Cell* 33: 729-740 (1983); Queen & Baltimore, *Cell* 33: 741-748 (1983)), neuron-specific promoters (e.g., the neurofilament promoter; Byrne & Ruddle, *Proc. Natl. Acad. Sci. USA* 86: 5473-5477 (1989)), pancreas-specific promoters (Edlund *et al.*, *Science* 230: 912-916 (1985)), and mammary gland-specific promoters (e.g., milk whey promoter; U.S. Patent No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are sometimes utilized, for example, the murine hox promoters (Kessel & Gruss, *Science* 249: 374-379 (1990)) and the α -fetoprotein promoter (Campes & Tilghman, *Genes Dev.* 3: 537-546 (1989)).

[0050] A *ADAMTS2* nucleic acid also may be cloned into an expression vector in an antisense orientation. Regulatory sequences (e.g., viral promoters and/or enhancers) operatively linked to a *ADAMTS2* nucleic acid cloned in the antisense orientation can be chosen for directing constitutive, tissue specific or cell type specific expression of antisense RNA in a variety of cell types. Antisense expression vectors can be in the form of a recombinant plasmid, phagemid or attenuated virus. For a discussion of the regulation of gene expression using antisense genes see, e.g., Weintraub *et al.*, Antisense RNA as a molecular tool for genetic analysis, *Reviews - Trends in Genetics*, Vol. 1(1) (1986).

[0051] Also provided herein are host cells that include a *ADAMTS2* nucleotide sequence within a recombinant expression vector or a fragment of such a nucleotide sequence which facilitate homologous recombination into a specific site of the host cell genome. The terms "host cell" and "recombinant host cell" are used interchangeably herein. Such terms refer not only to the particular subject cell but rather also to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein. A host cell can be any prokaryotic or eukaryotic cell. For example, a target polypeptide can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to those skilled in the art.

[0052] Vectors can be introduced into host cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (e.g., DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, transduction/infection, DEAE-dextran-mediated transfection, lipofection, or electroporation.

[0053] A host cell provided herein can be used to produce (*i.e.*, express) a target polypeptide or a substantially identical polypeptide thereof. Accordingly, further provided are methods for producing a target polypeptide using host cells described herein. In one embodiment, the method includes culturing host cells into which a recombinant expression vector encoding a target polypeptide has been introduced in a suitable medium such that a target polypeptide is produced. In another embodiment, the method further includes isolating a target polypeptide from the medium or the host cell.

[0054] Also provided are cells or purified preparations of cells which include a *ADAMTS2* transgene, or which otherwise misexpress target polypeptide. Cell preparations can consist of human or non-human cells, *e.g.*, rodent cells, *e.g.*, mouse or rat cells, rabbit cells, or pig cells. In preferred embodiments, the cell or cells include a *ADAMTS2* transgene (*e.g.*, a heterologous form of a *ADAMTS2* gene, such as a human gene expressed in non-human cells). The transgene can be misexpressed, *e.g.*, overexpressed or underexpressed. In other preferred embodiments, the cell or cells include a gene which misexpress an endogenous target polypeptide (*e.g.*, expression of a gene is disrupted, also known as a knockout). Such cells can serve as a model for studying disorders which are related to mutated or misexpressed alleles or for use in drug screening. Also provided are human cells (*e.g.*, a hematopoietic stem cells) transfected with a *ADAMTS2* nucleic acid.

[0055] Also provided are cells or a purified preparation thereof (*e.g.*, human cells) in which an endogenous *ADAMTS2* nucleic acid is under the control of a regulatory sequence that does not normally control the expression of the endogenous gene. The expression characteristics of an endogenous gene within a cell (*e.g.*, a cell line or microorganism) can be modified by inserting a heterologous DNA regulatory element into the genome of the cell such that the inserted regulatory element is operably linked to the corresponding endogenous gene. For example, an endogenous corresponding gene (*e.g.*, a gene which is “transcriptionally silent,” not normally expressed, or expressed only at very low levels) may be activated by inserting a regulatory element which is capable of promoting the expression of a normally expressed gene product in that cell. Techniques such as targeted homologous recombinations, can be used to insert the heterologous DNA as described in, *e.g.*, Chappel, US 5,272,071; WO 91/06667, published on May 16, 1991.

Transgenic Animals

[0056] Non-human transgenic animals that express a heterologous target polypeptide (*e.g.*, expressed from a *ADAMTS2* nucleic acid or substantially identical sequence thereof) can be generated. Such animals are useful for studying the function and/or activity of a target polypeptide and for identifying and/or evaluating modulators of the activity of *ADAMTS2* nucleic acids and encoded polypeptides. As used herein, a “transgenic animal” is a non-human animal such as a mammal (*e.g.*, a non-human primate such as chimpanzee, baboon, or macaque; an ungulate such as an equine, bovine, or

caprine; or a rodent such as a rat, a mouse, or an Israeli sand rat), a bird (e.g., a chicken or a turkey), an amphibian (e.g., a frog, salamander, or newt), or an insect (e.g., *Drosophila melanogaster*), in which one or more of the cells of the animal includes a transgene. A transgene is exogenous DNA or a rearrangement (e.g., a deletion of endogenous chromosomal DNA) that is often integrated into or occurs in the genome of cells in a transgenic animal. A transgene can direct expression of an encoded gene product in one or more cell types or tissues of the transgenic animal, and other transgenes can reduce expression (e.g., a knockout). Thus, a transgenic animal can be one in which an endogenous nucleic acid homologous to a *ADAMTS2* nucleic acid has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal (e.g., an embryonic cell of the animal) prior to development of the animal.

[0057] Intronic sequences and polyadenylation signals can also be included in the transgene to increase expression efficiency of the transgene. One or more tissue-specific regulatory sequences can be operably linked to a *ADAMTS2* nucleotide sequence to direct expression of an encoded polypeptide to particular cells. A transgenic founder animal can be identified based upon the presence of a *ADAMTS2* nucleotide sequence in its genome and/or expression of encoded mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a *ADAMTS2* nucleotide sequence can further be bred to other transgenic animals carrying other transgenes.

[0058] Target polypeptides can be expressed in transgenic animals or plants by introducing, for example, a *ADAMTS2* nucleic acid into the genome of an animal that encodes the target polypeptide. In preferred embodiments the nucleic acid is placed under the control of a tissue specific promoter, e.g., a milk or egg specific promoter, and recovered from the milk or eggs produced by the animal. Also included is a population of cells from a transgenic animal.

Target Polypeptides

[0059] Also featured herein are isolated target polypeptides, which are encoded by a *ADAMTS2* nucleotide sequence (e.g., SEQ ID NO: 1-3), or a substantially identical nucleotide sequence thereof. *ADAMTS2* exists in two forms, a "long" form comprising a molecule approximately 130 kDa in length (e.g., SEQ ID NO: 4), and a "short" form comprising a molecule approximately 70 kDa in length (e.g., SEQ ID NO: 5). Thus featured herein are isolated *ADAMTS2* polypeptides, which include long and short isoforms, and substantially identical polypeptides thereof. An *ADAMTS2* polypeptide is a polypeptide encoded by an *ADAMTS2* nucleic acid, where one nucleic acid can encode one or more different polypeptides. The term "polypeptide" as used herein includes proteins and peptides. An "isolated" or "purified" polypeptide or protein is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the protein is derived, or substantially free from chemical

precursors or other chemicals when chemically synthesized. In one embodiment, the language "substantially free" means preparation of a target polypeptide having less than about 30%, 20%, 10% and more preferably 5% (by dry weight), of non-target polypeptide (also referred to herein as a "contaminating protein"), or of chemical precursors or non-target chemicals. When the target polypeptide or a biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, specifically, where culture medium represents less than about 20%, sometimes less than about 10%, and often less than about 5% of the volume of the polypeptide preparation. Isolated or purified target polypeptide preparations are sometimes 0.01 milligrams or more or 0.1 milligrams or more, and often 1.0 milligrams or more and 10 milligrams or more in dry weight.

[0060] Further included herein are target polypeptide fragments. The polypeptide fragment may be a domain or part of a domain of a target polypeptide. The polypeptide fragment may have increased, decreased or unexpected biological activity. The polypeptide fragment is often 50 or fewer, 100 or fewer, or 200 or fewer amino acids in length, and is sometimes 300, 400, 500, 600, 700, or 900 or fewer amino acids in length. Thus, featured herein are *ADAMTS2* polypeptides and biologically active or antigenic fragments thereof useful as reagents or targets in assays applicable to treatment and diagnosis of osteoarthritis. In another embodiment, provided herein are *ADAMTS2* polypeptides having a *ADAMTS2* activity (e.g., a zinc binding activity, a metalloprotease activity, a procollagen II processing or synthesis activity, or a collagen II synthesis activity in vitro or in vivo). In certain embodiments, the polypeptides are *ADAMTS2* proteins including at least one propeptide domain, at least one metalloproteinase domain, at least one disintegrin-like domain, at least one, two, three, and often four thrombospondin domains, and sometimes having a *ADAMTS2* activity, e.g., a *ADAMTS2* activity as described herein. *ADAMTS2* polypeptides and fragments thereof often have biological activity, such as excising the N-propeptide of type II procollagens. Methods for monitoring and quantifying this biological activity are known (e.g., Colige et al., J. Biol. Chem. 270: 16724-16730 (1995)).

[0061] Human *ADAMTS2* protein (SEQ ID NO: 4-5) includes a signal sequence of about 29 amino acids (from amino acid 1 to about amino acid 29 of SEQ ID NO: 4-5). The *ADAMTS2* protein without the signal sequence can be approximately 1182 amino acid residues in length (from about amino acid 30 to amino acid 1211 of SEQ ID NO: 4) or approximately 485 amino acid residues in length (from about amino acid 30 to amino acid 514 of SEQ ID NO: 5). Human *ADAMTS2* protein includes a "pro" region homologous to the reproxin family propeptide, which is typically post-translationally cleaved upon conversion of the inactive (or pro-domain containing) protein to the catalytically active metalloprotease. The prodomain region of human *ADAMTS2* protein corresponds to about amino acids 30 to 251, 30 to 252, 30 to 253, 30 to 254, 30 to 255, 30 to 256, 30 to 257, 30 to 258 or 30 to 259 of SEQ ID NO: 4-5, where it is understood that the active form of *ADAMTS2* does not contain the propeptide domain.

[0062] Upon cleavage, catalytically active mature protein can be approximately 960, 959, 958, 957, 956, 955, 954, 953 or 952 amino acids in length (from about amino acid 252, 253, 254, 255, 256, 257, 258, 259 or 260 to amino acid 1211 of SEQ ID NO: 4) or approximately 261, 260, 259, 258, 257, 256, 255, 254 or 253 amino acid residues in length (from about amino acid 252, 253, 254, 255, 256, 257, 258, 259 or 260 to amino acid 514 of SEQ ID NO: 5).

[0063] Human *ADAMTS2* contains the following regions or other structural features: a signal sequence at about amino acids 1-29 of SEQ ID NO: 4-5; a reprolysin family propeptide domain located at about amino acid residues 30 to 251, 30 to 252, 30 to 253, 30 to 254, 30 to 255, 30 to 256, 30 to 257, 30 to 258 or 30 to 259 of SEQ ID NO: 4-5; a zinc-metalloprotease catalytic domain at about amino acids 251 to 479, 252 to 479, 253 to 479, 254 to 479, 255 to 479, 256 to 479, 257 to 479, 258 to 479 or 259 to 479 of SEQ ID NO: 4-5; a disintegrin domain at about amino acids 480 to 560 of SEQ ID NO: 4; a cysteine-rich domain at about amino acids 618 to 722 of SEQ ID NO: 4; four thrombospondin motifs-2 motifs at about amino acids 561 to 616, 854 to 912, 914 to 971, and 975 to 1029 of SEQ ID NO: 4; and eight N-glycosylation sites located at about amino acids 112, 251, 949, 993, 1031, 1098, 1145, and 1150 of SEQ ID NO: 4.

[0064] In other embodiments, provided are methods of increasing the synthesis of procollagen II comprising providing or administering to individuals in need of increasing levels of type II collagen the pharmaceutical or physiologically acceptable composition comprising active human *ADAMTS2* protein or fragment thereof, where *ADAMTS2* polypeptide fragments having activity are selected from amino acids 252-1211, 253-1211, 254-1211, 255-1211, 256-1211, 257-1211, 258-1211, 259-1211 or 260-1211 of SEQ ID NO: 4, where it is understood that the active form of *ADAMTS2* does not contain the propeptide domain.

[0065] Substantially identical target polypeptides may depart from the amino acid sequences of target polypeptides in different manners. For example, conservative amino acid modifications may be introduced at one or more positions in the amino acid sequences of target polypeptides. A “conservative amino acid substitution” is one in which the amino acid is replaced by another amino acid having a similar structure and/or chemical function. Families of amino acid residues having similar structures and functions are well known. These families include amino acids with basic side chains (e.g., lysine, arginine, histidine), acidic side chains (e.g., aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine). Also, essential and non-essential amino acids may be replaced. A “non-essential” amino acid is one that can be altered without abolishing or substantially altering the biological function of a target polypeptide, whereas altering an “essential” amino acid abolishes or substantially alters the

biological function of a target polypeptide. Amino acids that are conserved among target polypeptides are typically essential amino acids. In certain embodiments, the polypeptide includes one or more non-synonymous polymorphic variants associated with osteoarthritis.

[0066] Also, target polypeptides may exist as chimeric or fusion polypeptides. As used herein, a target “chimeric polypeptide” or target “fusion polypeptide” includes a target polypeptide linked to a non-target polypeptide. A “non-target polypeptide” refers to a polypeptide having an amino acid sequence corresponding to a polypeptide which is not substantially identical to the target polypeptide, which includes, for example, a polypeptide that is different from the target polypeptide and derived from the same or a different organism. The target polypeptide in the fusion polypeptide can correspond to an entire or nearly entire target polypeptide or a fragment thereof. The non-target polypeptide can be fused to the N-terminus or C-terminus of the target polypeptide.

[0067] Fusion polypeptides can include a moiety having high affinity for a ligand. For example, the fusion polypeptide can be a GST-target fusion polypeptide in which the target sequences are fused to the C-terminus of the GST sequences, or a polyhistidine-target fusion polypeptide in which the target polypeptide is fused at the N- or C-terminus to a string of histidine residues. Such fusion polypeptides can facilitate purification of recombinant target polypeptide. Expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide), and a nucleotide sequence in SEQ ID NO: 1-3, or a substantially identical nucleotide sequence thereof, can be cloned into an expression vector such that the fusion moiety is linked in-frame to the target polypeptide. Further, the fusion polypeptide can be a target polypeptide containing a heterologous signal sequence at its N-terminus. In certain host cells (*e.g.*, mammalian host cells), expression, secretion, cellular internalization, and cellular localization of a target polypeptide can be increased through use of a heterologous signal sequence. Fusion polypeptides can also include all or a part of a serum polypeptide (*e.g.*, an IgG constant region or human serum albumin).

[0068] Target polypeptides can be incorporated into pharmaceutical compositions and administered to a subject *in vivo*. Administration of these target polypeptides can be used to affect the bioavailability of a substrate of the target polypeptide and may effectively increase target polypeptide biological activity in a cell. Target fusion polypeptides may be useful therapeutically for the treatment of disorders caused by, for example, (i) aberrant modification or mutation of a gene encoding a target polypeptide; (ii) mis-regulation of the gene encoding the target polypeptide; and (iii) aberrant post-translational modification of a target polypeptide. Also, target polypeptides can be used as immunogens to produce anti-target antibodies in a subject, to purify target polypeptide ligands or binding partners, and in screening assays to identify molecules which inhibit or enhance the interaction of a target polypeptide with a substrate.

[0069] In addition, polypeptides can be chemically synthesized using techniques known in the art (See, *e.g.*, Creighton, 1983 Proteins. New York, N.Y.: W. H. Freeman and Company; and Hunkapiller et

al., (1984) Nature July 12 -18;310(5973):105-11). For example, a relative short fragment can be synthesized by use of a peptide synthesizer. Furthermore, if desired, non-classical amino acids or chemical amino acid analogs can be introduced as a substitution or addition into the fragment sequence. Non-classical amino acids include, but are not limited to, to the D-isomers of the common amino acids, 2,4-diaminobutyric acid, a-amino isobutyric acid, 4-aminobutyric acid, Abu, 2-amino butyric acid, g-Abu, e-Ahx, 6-amino hexanoic acid, Aib, 2-amino isobutyric acid, 3-amino propionic acid, ornithine, norleucine, norvaline, hydroxyproline, sarcosine, citrulline, homocitrulline, cysteic acid, t-butylglycine, t-butylalanine, phenylglycine, cyclohexylalanine, b-alanine, fluoroamino acids, designer amino acids such as b-methyl amino acids, Ca-methyl amino acids, Na-methyl amino acids, and amino acid analogs in general. Furthermore, the amino acid can be D (dextrorotary) or L (levorotary).

[0070] Polypeptides and polypeptide fragments sometimes are differentially modified during or after translation, *e.g.*, by glycosylation, acetylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to an antibody molecule or other cellular ligand, etc. Any of numerous chemical modifications may be carried out by known techniques, including but not limited, to specific chemical cleavage by cyanogen bromide, trypsin, chymotrypsin, papain, V8 protease, NaBH4; acetylation, formylation, oxidation, reduction; metabolic synthesis in the presence of tunicamycin; and the like. Additional post-translational modifications include, for example, N-linked or O-linked carbohydrate chains, processing of N-terminal or C-terminal ends), attachment of chemical moieties to the amino acid backbone, chemical modifications of N-linked or O-linked carbohydrate chains, and addition or deletion of an N-terminal methionine residue as a result of prokaryotic host cell expression. The polypeptide fragments may also be modified with a detectable label, such as an enzymatic, fluorescent, isotopic or affinity label to allow for detection and isolation of the polypeptide.

[0071] Also provided are chemically modified derivatives of polypeptides that can provide additional advantages such as increased solubility, stability and circulating time of the polypeptide, or decreased immunogenicity (*see e.g.*, U.S. Pat. No: 4,179,337). The chemical moieties for derivitization may be selected from water soluble polymers such as polyethylene glycol, ethylene glycol/propylene glycol copolymers, carboxymethylcellulose, dextran, polyvinyl alcohol and the like. The polypeptides may be modified at random positions within the molecule, or at predetermined positions within the molecule and may include one, two, three or more attached chemical moieties.

[0072] The polymer may be of any molecular weight, and may be branched or unbranched. For polyethylene glycol, the preferred molecular weight is between about 1 kDa and about 100 kDa (the term "about" indicating that in preparations of polyethylene glycol, some molecules will weigh more, some less, than the stated molecular weight) for ease in handling and manufacturing. Other sizes may be used, depending on the desired therapeutic profile (*e.g.*, the duration of sustained release desired, the effects, if

any on biological activity, the ease in handling, the degree or lack of antigenicity and other known effects of the polyethylene glycol to a therapeutic protein or analog).

[0073] The polymers should be attached to the polypeptide with consideration of effects on functional or antigenic domains of the polypeptide. There are a number of attachment methods available to those skilled in the art (e.g., EP 0 401 384 (coupling PEG to G-CSF) and Malik et al. (1992) Exp Hematol. September;20(8):1028-35 (pegylation of GM-CSF using tresyl chloride)). For example, polyethylene glycol may be covalently bound through amino acid residues via a reactive group, such as a free amino or carboxyl group. Reactive groups are those to which an activated polyethylene glycol molecule may be bound. The amino acid residues having a free amino group may include lysine residues and the N-terminal amino acid residues; those having a free carboxyl group may include aspartic acid residues, glutamic acid residues and the C-terminal amino acid residue. Sulfhydryl groups may also be used as a reactive group for attaching the polyethylene glycol molecules. For therapeutic purposes, the attachment sometimes is at an amino group, such as attachment at the N-terminus or lysine group.

[0074] Proteins can be chemically modified at the N-terminus. Using polyethylene glycol as an illustration of such a composition, one may select from a variety of polyethylene glycol molecules (by molecular weight, branching, and the like), the proportion of polyethylene glycol molecules to protein (polypeptide) molecules in the reaction mix, the type of pegylation reaction to be performed, and the method of obtaining the selected N-terminally pegylated protein. The method of obtaining the N-terminally pegylated preparation (i.e., separating this moiety from other monopegylated moieties if necessary) may be by purification of the N-terminally pegylated material from a population of pegylated protein molecules. Selective proteins chemically modified at the N-terminus may be accomplished by reductive alkylation, which exploits differential reactivity of different types of primary amino groups (lysine versus the N-terminal) available for derivatization in a particular protein. Under the appropriate reaction conditions, substantially selective derivatization of the protein at the N-terminus with a carbonyl group containing polymer is achieved.

Substantially Identical Nucleic Acids and Polypeptides

[0075] Nucleotide sequences and polypeptide sequences that are substantially identical to a *ADAMTS2* nucleotide sequence and the target polypeptide sequences encoded by those nucleotide sequences, respectively, are included herein. The term "substantially identical" as used herein refers to two or more nucleic acids or polypeptides sharing one or more identical nucleotide sequences or polypeptide sequences, respectively. Included are nucleotide sequences or polypeptide sequences that are 55% or more, 60% or more, 65% or more, 70% or more, 75% or more, 80% or more, 85% or more, 90% or more, 95% or more (each often within a 1%, 2%, 3% or 4% variability) identical to a *ADAMTS2* nucleotide sequence or the encoded target polypeptide amino acid sequences. One test for determining

whether two nucleic acids are substantially identical is to determine the percent of identical nucleotide sequences or polypeptide sequences shared between the nucleic acids or polypeptides.

[0076] Calculations of sequence identity are often performed as follows. Sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-homologous sequences can be disregarded for comparison purposes). The length of a reference sequence aligned for comparison purposes is sometimes 30% or more, 40% or more, 50% or more, often 60% or more, and more often 70% or more, 80% or more, 90% or more, or 100% of the length of the reference sequence. The nucleotides or amino acids at corresponding nucleotide or polypeptide positions, respectively, are then compared among the two sequences. When a position in the first sequence is occupied by the same nucleotide or amino acid as the corresponding position in the second sequence, the nucleotides or amino acids are deemed to be identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, introduced for optimal alignment of the two sequences.

[0077] Comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm. Percent identity between two amino acid or nucleotide sequences can be determined using the algorithm of Meyers & Miller, *CABIOS* 4: 11-17 (1989), which has been incorporated into the ALIGN program (version 2.0), using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4. Also, percent identity between two amino acid sequences can be determined using the Needleman & Wunsch, *J. Mol. Biol.* 48: 444-453 (1970) algorithm which has been incorporated into the GAP program in the GCG software package (available at the http address www.gcg.com), using either a Blossum 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. Percent identity between two nucleotide sequences can be determined using the GAP program in the GCG software package (available at http address www.gcg.com), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. A set of parameters often used is a Blossum 62 scoring matrix with a gap open penalty of 12, a gap extend penalty of 4, and a frameshift gap penalty of 5.

[0078] Another manner for determining if two nucleic acids are substantially identical is to assess whether a polynucleotide homologous to one nucleic acid will hybridize to the other nucleic acid under stringent conditions. As use herein, the term "stringent conditions" refers to conditions for hybridization and washing. Stringent conditions are known to those skilled in the art and can be found in *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y., 6.3.1-6.3.6 (1989). Aqueous and non-aqueous methods are described in that reference and either can be used. An example of stringent hybridization conditions is hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C,

followed by one or more washes in 0.2X SSC, 0.1% SDS at 50°C. Another example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 55°C. A further example of stringent hybridization conditions is hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 60°C. Often, stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 65°C. More often, stringency conditions are 0.5M sodium phosphate, 7% SDS at 65°C, followed by one or more washes at 0.2X SSC, 1% SDS at 65°C.

[0079] An example of a substantially identical nucleotide sequence to a nucleotide sequence in SEQ ID NO: 1-3 is one that has a different nucleotide sequence but still encodes the same polypeptide sequence encoded by the nucleotide sequence in SEQ ID NO: 1-3. Another example is a nucleotide sequence that encodes a polypeptide having a polypeptide sequence that is more than 70% or more identical to, sometimes more than 75% or more, 80% or more, or 85% or more identical to, and often more than 90% or more and 95% or more identical to a polypeptide sequence encoded by a nucleotide sequence in SEQ ID NO: 1-3.

[0080] Nucleotide sequences in SEQ ID NO: 1-3 and amino acid sequences of encoded polypeptides can be used as “query sequences” to perform a search against public databases to identify other family members or related sequences, for example. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul *et al.*, *J. Mol. Biol.* 215: 403-10 (1990). BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to nucleotide sequences in SEQ ID NO: 1-3. BLAST polypeptide searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to polypeptides encoded by the nucleotide sequences of SEQ ID NO: 1-3. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul *et al.*, *Nucleic Acids Res.* 25(17): 3389-3402 (1997). When utilizing BLAST and Gapped BLAST programs, default parameters of the respective programs (*e.g.*, XBLAST and NBLAST) can be used (*see* the [http address www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)).

[0081] A nucleic acid that is substantially identical to a nucleotide sequence in SEQ ID NO: 1-3 may include polymorphic sites at positions equivalent to those described herein when the sequences are aligned. For example, using the alignment procedures described herein, SNPs in a sequence substantially identical to a sequence in SEQ ID NO: 1-3 can be identified at nucleotide positions that match (*i.e.*, align) with nucleotides at SNP positions in each nucleotide sequence in SEQ ID NO: 1-3. Also, where a polymorphic variation results in an insertion or deletion, insertion or deletion of a nucleotide sequence

from a reference sequence can change the relative positions of other polymorphic sites in the nucleotide sequence.

[0082] Substantially identical nucleotide and polypeptide sequences include those that are naturally occurring, such as allelic variants (same locus), splice variants, homologs (different locus), and orthologs (different organism) or can be non-naturally occurring. Non-naturally occurring variants can be generated by mutagenesis techniques, including those applied to polynucleotides, cells, or organisms. The variants can contain nucleotide substitutions, deletions, inversions and insertions. Variation can occur in either or both the coding and non-coding regions. The variations can produce both conservative and non-conservative amino acid substitutions (as compared in the encoded product). Orthologs, homologs, allelic variants, and splice variants can be identified using methods known in the art. These variants normally comprise a nucleotide sequence encoding a polypeptide that is 50% or more, about 55% or more, often about 70-75% or more or about 80-85% or more, and sometimes about 90-95% or more identical to the amino acid sequences of target polypeptides or a fragment thereof. Such nucleic acid molecules can readily be identified as being able to hybridize under stringent conditions to a nucleotide sequence in SEQ ID NO: 1-3 or a fragment of this sequence. Nucleic acid molecules corresponding to orthologs, homologs, and allelic variants of a nucleotide sequence in SEQ ID NO: 1-3 can further be identified by mapping the sequence to the same chromosome or locus as the nucleotide sequence in SEQ ID NO: 1-3.

[0083] Also, substantially identical nucleotide sequences may include codons that are altered with respect to the naturally occurring sequence for enhancing expression of a target polypeptide in a particular expression system. For example, the nucleic acid can be one in which one or more codons are altered, and often 10% or more or 20% or more of the codons are altered for optimized expression in bacteria (*e.g.*, *E. coli*), yeast (*e.g.*, *S. cerevisiae*), human (*e.g.*, 293 cells), insect, or rodent (*e.g.*, hamster) cells.

Methods for Identifying Risk of Osteoarthritis

[0084] Methods for prognosing and diagnosing osteoarthritis are included herein. These methods include detecting the presence or absence of one or more polymorphic variations in a nucleotide sequence associated with osteoarthritis, such as variants in or around the loci set forth herein, or a substantially identical sequence thereof, in a sample from a subject, where the presence of a polymorphic variant described herein is indicative of a risk of osteoarthritis. Determining a risk of osteoarthritis sometimes refers to determining whether an individual is at an increased risk of osteoarthritis (*e.g.*, intermediate risk or higher risk).

[0085] Thus, featured herein is a method for identifying a subject who is at risk of osteoarthritis, which comprises detecting an aberration associated with osteoarthritis in a nucleic acid sample from the

subject. An embodiment is a method for detecting a risk of osteoarthritis in a subject, which comprises detecting the presence or absence of a polymorphic variation associated with osteoarthritis at a polymorphic site in a nucleotide sequence in a nucleic acid sample from a subject, where the nucleotide sequence comprises a polynucleotide sequence selected from the group consisting of: (a) a nucleotide sequence of SEQ ID NO: 1-3; (b) a nucleotide sequence which encodes a polypeptide consisting of an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1-3; (c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1-3, or a nucleotide sequence about 90% or more identical to a nucleotide sequence of SEQ ID NO: 1-3; and (d) a fragment of a nucleotide sequence of (a), (b), or (c) comprising the polymorphic site; whereby the presence of the polymorphic variation is indicative of a predisposition to osteoarthritis in the subject. In some embodiments, a polymorphic variation at position 733 of SEQ ID NO: 2-3 may be detected (e.g., a guanine in these sequences or a cytosine in a complementary sequence are associated with increased risk of osteoarthritis). In certain embodiments, polymorphic variants at the positions described herein are detected for determining a risk of osteoarthritis, and polymorphic variants at positions in linkage disequilibrium with these positions are detected for determining a risk of osteoarthritis. As used herein, "SEQ ID NO: 1-3" refers to individual sequences in SEQ ID NO: 1, 2 or 3, each sequence being separately applicable to embodiments described herein.

[0086] Risk of osteoarthritis sometimes is expressed as a probability, such as an odds ratio, percentage, or risk factor. Risk often is based upon the presence or absence of one or more polymorphic variants described herein, and also may be based in part upon phenotypic traits of the individual being tested. Methods for calculating risk based upon patient data are well known (see, e.g., Agresti, *Categorical Data Analysis*, 2nd Ed. 2002. Wiley). Allelotyping and genotyping analyses may be carried out in populations other than those exemplified herein to enhance the predictive power of the prognostic method. These further analyses are executed in view of the exemplified procedures described herein, and may be based upon the same polymorphic variations or additional polymorphic variations.

[0087] In certain embodiments, determining the presence of a combination of two or more polymorphic variants associated with osteoarthritis in one or more genetic loci (e.g., one or more genes) of the sample is determined to identify, quantify and/or estimate, risk of osteoarthritis. The risk often is the probability of having or developing osteoarthritis. The risk sometimes is expressed as a relative risk with respect to a population average risk of osteoarthritis, and sometimes is expressed as a relative risk with respect to the lowest risk group. Such relative risk assessments often are based upon penetrance values determined by statistical methods, and are particularly useful to clinicians and insurance companies for assessing risk of osteoarthritis (e.g., a clinician can target appropriate detection, prevention and therapeutic regimens to a patient after determining the patient's risk of osteoarthritis, and an

insurance company can fine tune actuarial tables based upon population genotype assessments of osteoarthritis risk). Risk of osteoarthritis sometimes is expressed as an odds ratio, which is the odds of a particular person having a genotype has or will develop osteoarthritis with respect to another genotype group (e.g., the most disease protective genotype or population average). In related embodiments, the determination is utilized to identify a subject at risk of osteoarthritis. In an embodiment, two or more polymorphic variations are detected in two or more regions in human genomic DNA associated with increased risk of osteoarthritis, such as a locus containing a *ADAMTS2*, for example. In certain embodiments, 3 or more, or 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 or more polymorphic variants are detected in the sample. In specific embodiments, polymorphic variants are detected in a *ADAMTS2* region, for example. In certain embodiments, polymorphic variants are detected at other genetic loci (e.g., the polymorphic variants can be detected in *ADAMTS2* in addition to other loci or only in other loci), where the other loci include but are not limited to those described in concurrently-filed patent applications having attorney docket number 524593008700, 524593008800, 524593009000 or 524593009200, which is incorporated herein by reference in its entirety.

[0088] Results from prognostic tests may be combined with other test results to diagnose osteoarthritis. For example, prognostic results may be gathered, a patient sample may be ordered based on a determined predisposition to osteoarthritis, the patient sample is analyzed, and the results of the analysis may be utilized to diagnose osteoarthritis. Also osteoarthritis diagnostic method can be developed from studies used to generate prognostic methods in which populations are stratified into subpopulations having different progressions of osteoarthritis. In another embodiment, prognostic results may be gathered, a patient's risk factors for developing osteoarthritis (e.g., age, weight, race, diet) analyzed, and a patient sample may be ordered based on a determined predisposition to osteoarthritis.

[0089] The nucleic acid sample typically is isolated from a biological sample obtained from a subject. For example, nucleic acid can be isolated from blood, saliva, sputum, urine, cell scrapings, and biopsy tissue. The nucleic acid sample can be isolated from a biological sample using standard techniques, such as the technique described in Example 2. As used herein, the term "subject" refers primarily to humans but also refers to other mammals such as dogs, cats, and ungulates (e.g., cattle, sheep, and swine). Subjects also include avians (e.g., chickens and turkeys), reptiles, and fish (e.g., salmon), as embodiments described herein can be adapted to nucleic acid samples isolated from any of these organisms. The nucleic acid sample may be isolated from the subject and then directly utilized in a method for determining the presence of a polymorphic variant, or alternatively, the sample may be isolated and then stored (e.g., frozen) for a period of time before being subjected to analysis.

[0090] The presence or absence of a polymorphic variant is determined using one or both chromosomal complements represented in the nucleic acid sample. Determining the presence or absence of a polymorphic variant in both chromosomal complements represented in a nucleic acid sample from a

subject having a copy of each chromosome is useful for determining the zygosity of an individual for the polymorphic variant (*i.e.*, whether the individual is homozygous or heterozygous for the polymorphic variant). Any oligonucleotide-based diagnostic may be utilized to determine whether a sample includes the presence or absence of a polymorphic variant in a sample. For example, primer extension methods, ligase sequence determination methods (*e.g.*, U.S. Pat. Nos. 5,679,524 and 5,952,174, and WO 01/27326), mismatch sequence determination methods (*e.g.*, U.S. Pat. Nos. 5,851,770; 5,958,692; 6,110,684; and 6,183,958), microarray sequence determination methods, restriction fragment length polymorphism (RFLP), single strand conformation polymorphism detection (SSCP) (*e.g.*, U.S. Pat. Nos. 5,891,625 and 6,013,499), PCR-based assays (*e.g.*, TAQMAN® PCR System (Applied Biosystems)), and nucleotide sequencing methods may be used.

[0091] Oligonucleotide extension methods typically involve providing a pair of oligonucleotide primers in a polymerase chain reaction (PCR) or in other nucleic acid amplification methods for the purpose of amplifying a region from the nucleic acid sample that comprises the polymorphic variation. One oligonucleotide primer is complementary to a region 3' of the polymorphism and the other is complementary to a region 5' of the polymorphism. A PCR primer pair may be used in methods disclosed in U.S. Pat. Nos. 4,683,195; 4,683,202, 4,965,188; 5,656,493; 5,998,143; 6,140,054; WO 01/27327; and WO 01/27329 for example. PCR primer pairs may also be used in any commercially available machines that perform PCR, such as any of the GENEAMP® Systems available from Applied Biosystems. Also, those of ordinary skill in the art will be able to design oligonucleotide primers based upon a *ADAMTS2* nucleotide sequence using knowledge available in the art.

[0092] Also provided is an extension oligonucleotide that hybridizes to the amplified fragment adjacent to the polymorphic variation. As used herein, the term “adjacent” refers to the 3' end of the extension oligonucleotide being often 1 nucleotide from the 5' end of the polymorphic site, and sometimes 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides from the 5' end of the polymorphic site, in the nucleic acid when the extension oligonucleotide is hybridized to the nucleic acid. The extension oligonucleotide then is extended by one or more nucleotides, and the number and/or type of nucleotides that are added to the extension oligonucleotide determine whether the polymorphic variant is present. Oligonucleotide extension methods are disclosed, for example, in U.S. Pat. Nos. 4,656,127; 4,851,331; 5,679,524; 5,834,189; 5,876,934; 5,908,755; 5,912,118; 5,976,802; 5,981,186; 6,004,744; 6,013,431; 6,017,702; 6,046,005; 6,087,095; 6,210,891; and WO 01/20039. Oligonucleotide extension methods using mass spectrometry are described, for example, in U.S. Pat. Nos. 5,547,835; 5,605,798; 5,691,141; 5,849,542; 5,869,242; 5,928,906; 6,043,031; and 6,194,144, and a method often utilized is described herein in Example 2.

[0093] A microarray can be utilized for determining whether a polymorphic variant is present or absent in a nucleic acid sample. A microarray may include any oligonucleotides described herein, and

methods for making and using oligonucleotide microarrays suitable for diagnostic use are disclosed in U.S. Pat. Nos. 5,492,806; 5,525,464; 5,589,330; 5,695,940; 5,849,483; 6,018,041; 6,045,996; 6,136,541; 6,142,681; 6,156,501; 6,197,506; 6,223,127; 6,225,625; 6,229,911; 6,239,273; WO 00/52625; WO 01/25485; and WO 01/29259. The microarray typically comprises a solid support and the oligonucleotides may be linked to this solid support by covalent bonds or by non-covalent interactions. The oligonucleotides may also be linked to the solid support directly or by a spacer molecule. A microarray may comprise one or more oligonucleotides complementary to a polymorphic site set forth herein.

[0094] A kit also may be utilized for determining whether a polymorphic variant is present or absent in a nucleic acid sample. A kit often comprises one or more pairs of oligonucleotide primers useful for amplifying a fragment of a nucleotide sequence of SEQ ID NO: 1-3 or a substantially identical sequence thereof, where the fragment includes a polymorphic site. The kit sometimes comprises a polymerizing agent, for example, a thermostable nucleic acid polymerase such as one disclosed in U.S. Pat. Nos. 4,889,818 or 6,077,664. Also, the kit often comprises an elongation oligonucleotide that hybridizes to a *ADAMTS2* nucleotide sequence in a nucleic acid sample adjacent to the polymorphic site. Where the kit includes an elongation oligonucleotide, it also often comprises chain elongating nucleotides, such as dATP, dTTP, dGTP, dCTP, and dITP, including analogs of dATP, dTTP, dGTP, dCTP and dITP, provided that such analogs are substrates for a thermostable nucleic acid polymerase and can be incorporated into a nucleic acid chain elongated from the extension oligonucleotide. Along with chain elongating nucleotides would be one or more chain terminating nucleotides such as ddATP, ddTTP, ddGTP, ddCTP, and the like. In an embodiment, the kit comprises one or more oligonucleotide primer pairs, a polymerizing agent, chain elongating nucleotides, at least one elongation oligonucleotide, and one or more chain terminating nucleotides. Kits optionally include buffers, vials, microtiter plates, and instructions for use.

[0095] An individual identified as being at risk of osteoarthritis may be heterozygous or homozygous with respect to the allele associated with a higher risk of osteoarthritis. A subject homozygous for an allele associated with an increased risk of osteoarthritis is at a comparatively high risk of osteoarthritis, a subject heterozygous for an allele associated with an increased risk of osteoarthritis is at a comparatively intermediate risk of osteoarthritis, and a subject homozygous for an allele associated with a decreased risk of osteoarthritis is at a comparatively low risk of osteoarthritis. A genotype may be assessed for a complementary strand, such that the complementary nucleotide at a particular position is detected.

[0096] Also featured are methods for determining risk of osteoarthritis and/or identifying a subject at risk of osteoarthritis by contacting a polypeptide or protein encoded by a *ADAMTS2* nucleotide sequence from a subject with an antibody that specifically binds to an epitope associated with increased

risk of osteoarthritis in the polypeptide (e.g., an epitope comprising a valine at position 245 in an *IRILR1* polypeptide).

Applications of Prognostic and Diagnostic Results to Pharmacogenomic Methods

[0097] Pharmacogenomics is a discipline that involves tailoring a treatment for a subject according to the subject's genotype as a particular treatment regimen may exert a differential effect depending upon the subject's genotype. For example, based upon the outcome of a prognostic test described herein, a clinician or physician may target pertinent information and preventative or therapeutic treatments to a subject who would be benefited by the information or treatment and avoid directing such information and treatments to a subject who would not be benefited (e.g., the treatment has no therapeutic effect and/or the subject experiences adverse side effects).

[0098] The following is an example of a pharmacogenomic embodiment. A particular treatment regimen can exert a differential effect depending upon the subject's genotype. Where a candidate therapeutic exhibits a significant interaction with a major allele and a comparatively weak interaction with a minor allele (e.g., an order of magnitude or greater difference in the interaction), such a therapeutic typically would not be administered to a subject genotyped as being homozygous for the minor allele, and sometimes not administered to a subject genotyped as being heterozygous for the minor allele. In another example, where a candidate therapeutic is not significantly toxic when administered to subjects who are homozygous for a major allele but is comparatively toxic when administered to subjects heterozygous or homozygous for a minor allele, the candidate therapeutic is not typically administered to subjects who are genotyped as being heterozygous or homozygous with respect to the minor allele.

[0099] The methods described herein are applicable to pharmacogenomic methods for preventing, alleviating or treating osteoarthritis. For example, a nucleic acid sample from an individual may be subjected to a prognostic test described herein. Where one or more polymorphic variations associated with increased risk of osteoarthritis are identified in a subject, information for preventing or treating osteoarthritis and/or one or more osteoarthritis treatment regimens then may be prescribed to that subject.

[0100] In certain embodiments, a treatment or preventative regimen is specifically prescribed and/or administered to individuals who will most benefit from it based upon their risk of developing osteoarthritis assessed by the methods described herein. Thus, provided are methods for identifying a subject predisposed to osteoarthritis and then prescribing a therapeutic or preventative regimen to individuals identified as having a predisposition. Thus, certain embodiments are directed to a method for reducing osteoarthritis in a subject, which comprises: detecting the presence or absence of a polymorphic variant associated with osteoarthritis in a nucleotide sequence in a nucleic acid sample from a subject, where the nucleotide sequence comprises a polynucleotide sequence selected from the group consisting of: (a) a nucleotide sequence of SEQ ID NO: 1-3; (b) a nucleotide sequence which encodes a polypeptide

consisting of an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1-3; (c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1-3, or a nucleotide sequence about 90% or more identical to a nucleotide sequence of SEQ ID NO: 1-3; and (d) a fragment of a polynucleotide sequence of (a), (b), or (c); and prescribing or administering a treatment regimen to a subject from whom the sample originated where the presence of a polymorphic variation associated with osteoarthritis is detected in the nucleotide sequence. In these methods, predisposition results may be utilized in combination with other test results to diagnose osteoarthritis.

[0101] Certain preventative treatments often are prescribed to subjects having a predisposition to osteoarthritis and where the subject is diagnosed with osteoarthritis or is diagnosed as having symptoms indicative of an early stage of osteoarthritis. The treatment sometimes is preventative (e.g., is prescribed or administered to reduce the probability that osteoarthritis arises or progresses), sometimes is therapeutic, and sometimes delays, alleviates or halts the progression of osteoarthritis. Any known preventative or therapeutic treatment for alleviating or preventing the occurrence of osteoarthritis is prescribed and/or administered. For example, the treatment often is directed to decreasing pain and improving joint movement. Examples of OA treatments include exercises to keep joints flexible and improve muscle strength. Different medications to control pain, including corticosteroids and nonsteroidal anti-inflammatory drugs (NSAIDs, e.g., Voltaren); cyclooxygenase-2 (COX-2) inhibitors (e.g., Celebrex, Vioxx, Mobic, and Bextra); monoclonal antibodies (e.g., Remicade); tumor necrosis factor inhibitors (e.g., Enbrel); or injections of glucocorticoids, hyaluronic acid or chondroitin sulfate into joints that are inflamed and not responsive to NSAIDS. Orally administered chondroitin sulfate also may be used as a therapeutic, as it may increase hyaluronic acid levels and viscosity of synovial fluid, and decrease collagenase levels in synovial fluid. Also, glucosamine can serve as an OA therapeutic as delivering it into joints may inhibit enzymes involved in cartilage degradation and enhance the production of hyaluronic acid. For mild pain without inflammation, acetaminophen may be used. Other treatments include: heat/cold therapy for temporary pain relief; joint protection to prevent strain or stress on painful joints; surgery to relieve chronic pain in damaged joints; and weight control to prevent extra stress on weight-bearing joints.

[0102] As therapeutic approaches for treating osteoarthritis continue to evolve and improve, the goal of treatments for osteoarthritis related disorders is to intervene even before clinical signs first manifest. Thus, genetic markers associated with susceptibility to osteoarthritis prove useful for early diagnosis, prevention and treatment of osteoarthritis.

[0103] As osteoarthritis preventative and treatment information can be specifically targeted to subjects in need thereof (e.g., those at risk of developing osteoarthritis or those in an early stage of osteoarthritis), provided herein is a method for preventing or reducing the risk of developing

osteoarthritis in a subject, which comprises: (a) detecting the presence or absence of a polymorphic variation associated with osteoarthritis at a polymorphic site in a nucleotide sequence in a nucleic acid sample from a subject; (b) identifying a subject with a predisposition to osteoarthritis, whereby the presence of the polymorphic variation is indicative of a predisposition to osteoarthritis in the subject; and (c) if such a predisposition is identified, providing the subject with information about methods or products to prevent or reduce osteoarthritis or to delay the onset of osteoarthritis. Also provided is a method of targeting information or advertising to a subpopulation of a human population based on the subpopulation being genetically predisposed to a disease or condition, which comprises: (a) detecting the presence or absence of a polymorphic variation associated with osteoarthritis at a polymorphic site in a nucleotide sequence in a nucleic acid sample from a subject; (b) identifying the subpopulation of subjects in which the polymorphic variation is associated with osteoarthritis; and (c) providing information only to the subpopulation of subjects about a particular product which may be obtained and consumed or applied by the subject to help prevent or delay onset of the disease or condition.

[0104] Pharmacogenomics methods also may be used to analyze and predict a response to osteoarthritis treatment or a drug. For example, if pharmacogenomics analysis indicates a likelihood that an individual will respond positively to osteoarthritis treatment with a particular drug, the drug may be administered to the individual. Conversely, if the analysis indicates that an individual is likely to respond negatively to treatment with a particular drug, an alternative course of treatment may be prescribed. A negative response may be defined as either the absence of an efficacious response or the presence of toxic side effects. The response to a therapeutic treatment can be predicted in a background study in which subjects in any of the following populations are genotyped: a population that responds favorably to a treatment regimen, a population that does not respond significantly to a treatment regimen, and a population that responds adversely to a treatment regimen (*e.g.*, exhibits one or more side effects). These populations are provided as examples and other populations and subpopulations may be analyzed. Based upon the results of these analyses, a subject is genotyped to predict whether he or she will respond favorably to a treatment regimen, not respond significantly to a treatment regimen, or respond adversely to a treatment regimen.

[0105] The tests described herein also are applicable to clinical drug trials. One or more polymorphic variants indicative of response to an agent for treating osteoarthritis or to side effects to an agent for treating osteoarthritis may be identified using the methods described herein. Thereafter, potential participants in clinical trials of such an agent may be screened to identify those individuals most likely to respond favorably to the drug and exclude those likely to experience side effects. In that way, the effectiveness of drug treatment may be measured in individuals who respond positively to the drug, without lowering the measurement as a result of the inclusion of individuals who are unlikely to respond positively in the study and without risking undesirable safety problems.

[0106] Thus, another embodiment is a method of selecting an individual for inclusion in a clinical trial of a treatment or drug comprising the steps of: (a) obtaining a nucleic acid sample from an individual; (b) determining the identity of a polymorphic variation which is associated with a positive response to the treatment or the drug, or at least one polymorphic variation which is associated with a negative response to the treatment or the drug in the nucleic acid sample, and (c) including the individual in the clinical trial if the nucleic acid sample contains said polymorphic variation associated with a positive response to the treatment or the drug or if the nucleic acid sample lacks said polymorphic variation associated with a negative response to the treatment or the drug. In addition, the methods described herein for selecting an individual for inclusion in a clinical trial of a treatment or drug encompass methods with any further limitation described in this disclosure, or those following, specified alone or in any combination. The polymorphic variation may be in a sequence selected individually or in any combination from the group consisting of (i) a nucleotide sequence of SEQ ID NO: 1-3; (ii) a nucleotide sequence which encodes a polypeptide consisting of an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1-3; (iii) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1-3, or a nucleotide sequence about 90% or more identical to a nucleotide sequence of SEQ ID NO: 1-3; and (iv) a fragment of a polynucleotide sequence of (i), (ii), or (iii) comprising the polymorphic site. The including step (c) optionally comprises administering the drug or the treatment to the individual if the nucleic acid sample contains the polymorphic variation associated with a positive response to the treatment or the drug and the nucleic acid sample lacks said biallelic marker associated with a negative response to the treatment or the drug.

[0107] Also provided herein is a method of partnering between a diagnostic/prognostic testing provider and a provider of a consumable product, which comprises: (a) the diagnostic/prognostic testing provider detects the presence or absence of a polymorphic variation associated with osteoarthritis at a polymorphic site in a nucleotide sequence in a nucleic acid sample from a subject; (b) the diagnostic/prognostic testing provider identifies the subpopulation of subjects in which the polymorphic variation is associated with osteoarthritis; (c) the diagnostic/prognostic testing provider forwards information to the subpopulation of subjects about a particular product which may be obtained and consumed or applied by the subject to help prevent or delay onset of the disease or condition; and (d) the provider of a consumable product forwards to the diagnostic test provider a fee every time the diagnostic/prognostic test provider forwards information to the subject as set forth in step (c) above.

Compositions Comprising Osteoarthritis-Directed Molecules

[0108] Featured herein is a composition comprising a cell from a subject having osteoarthritis or at risk of osteoarthritis and one or more molecules specifically directed and targeted to a nucleic acid

comprising a *ADAMTS2* nucleotide sequence or amino acid sequence. Such directed molecules include, but are not limited to, a compound that binds to a *ADAMTS2* nucleotide sequence or amino acid sequence referenced herein; a RNAi or siRNA molecule having a strand complementary or substantially complementary to a *ADAMTS2* nucleotide sequence (e.g., hybridizes to a *ADAMTS2* nucleotide sequence under conditions of high stringency); an antisense nucleic acid complementary or substantially complementary to an RNA encoded by a *ADAMTS2* nucleotide sequence (e.g., hybridizes to a *ADAMTS2* nucleotide sequence under conditions of high stringency); a ribozyme that hybridizes to a *ADAMTS2* nucleotide sequence (e.g., hybridizes to a *ADAMTS2* nucleotide sequence under conditions of high stringency); a nucleic acid aptamer that specifically binds a polypeptide encoded by *ADAMTS2* nucleotide sequence; and an antibody that specifically binds to a polypeptide encoded by *ADAMTS2* nucleotide sequence or binds to a nucleic acid having such a nucleotide sequence. In specific embodiments, the osteoarthritis directed molecule interacts with a nucleic acid or polypeptide variant associated with osteoarthritis, such as variants referenced herein. In other embodiments, the osteoarthritis directed molecule interacts with a polypeptide involved in a signal pathway of a polypeptide encoded by a *ADAMTS2* nucleotide sequence, or a nucleic acid comprising such a nucleotide sequence.

[0109] Compositions sometimes include an adjuvant known to stimulate an immune response, and in certain embodiments, an adjuvant that stimulates a T-cell lymphocyte response. Adjuvants are known, including but not limited to an aluminum adjuvant (e.g., aluminum hydroxide); a cytokine adjuvant or adjuvant that stimulates a cytokine response (e.g., interleukin (IL)-12 and/or gamma-interferon cytokines); a Freund-type mineral oil adjuvant emulsion (e.g., Freund's complete or incomplete adjuvant); a synthetic lipoid compound; a copolymer adjuvant (e.g., TitreMax); a saponin; Quil A; a liposome; an oil-in-water emulsion (e.g., an emulsion stabilized by Tween 80 and pluronic polyoxyethylene/polyoxypropylene block copolymer (Syntex Adjuvant Formulation); TitreMax; detoxified endotoxin (MPL) and mycobacterial cell wall components (TDW, CWS) in 2% squalene (Ribi Adjuvant System)); a muramyl dipeptide; an immune-stimulating complex (ISCOM, e.g., an Ag-modified saponin/cholesterol micelle that forms stable cage-like structure); an aqueous phase adjuvant that does not have a depot effect (e.g., Gerbu adjuvant); a carbohydrate polymer (e.g., AdjuPrime); L-tyrosine; a manide-oleate compound (e.g., Montanide); an ethylene-vinyl acetate copolymer (e.g., Elvax 40W1,2); or lipid A, for example. Such compositions are useful for generating an immune response against osteoarthritis directed molecule (e.g., an HLA-binding subsequence within a polypeptide encoded by a *ADAMTS2* nucleotide sequence). In such methods, a peptide having an amino acid subsequence of a polypeptide encoded by a *ADAMTS2* nucleotide sequence is delivered to a subject, where the subsequence binds to an HLA molecule and induces a CTL lymphocyte response. The peptide sometimes is delivered to the subject as an isolated peptide or as a minigene in a plasmid that encodes the

peptide. Methods for identifying HLA-binding subsequences in such polypeptides are known (see e.g., publication WO02/20616 and PCT application US98/01373 for methods of identifying such sequences).

[0110] The cell may be in a group of cells cultured *in vitro* or in a tissue maintained *in vitro* or present in an animal *in vivo* (e.g., a rat, mouse, ape or human). In certain embodiments, a composition comprises a component from a cell such as a nucleic acid molecule (e.g., genomic DNA), a protein mixture or isolated protein, for example. The aforementioned compositions have utility in diagnostic, prognostic and pharmacogenomic methods described previously and in therapeutics described hereafter. Certain osteoarthritis directed molecules are described in greater detail below.

Compounds

[0111] Compounds can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; peptoid libraries (libraries of molecules having the functionalities of peptides, but with a novel, non-peptide backbone which are resistant to enzymatic degradation but which nevertheless remain bioactive (see, e.g., Zuckermann et al., J. Med. Chem. 37: 2678-85 (1994)); spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; “one-bead one-compound” library methods; and synthetic library methods using affinity chromatography selection. Biological library and peptoid library approaches are typically limited to peptide libraries, while the other approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam, Anticancer Drug Des. 12: 145, (1997)). Examples of methods for synthesizing molecular libraries are described, for example, in DeWitt et al., Proc. Natl. Acad. Sci. U.S.A. 90: 6909 (1993); Erb et al., Proc. Natl. Acad. Sci. USA 91: 11422 (1994); Zuckermann et al., J. Med. Chem. 37: 2678 (1994); Cho et al., Science 261: 1303 (1993); Carell et al., Angew. Chem. Int. Ed. Engl. 33: 2059 (1994); Carell et al., Angew. Chem. Int. Ed. Engl. 33: 2061 (1994); and in Gallop et al., J. Med. Chem. 37: 1233 (1994).

[0112] Libraries of compounds may be presented in solution (e.g., Houghten, Biotechniques 13: 412-421 (1992)), or on beads (Lam, Nature 354: 82-84 (1991)), chips (Fodor, Nature 364: 555-556 (1993)), bacteria or spores (Ladner, United States Patent No. 5,223,409), plasmids (Cull et al., Proc. Natl. Acad. Sci. USA 89: 1865-1869 (1992)) or on phage (Scott and Smith, Science 249: 386-390 (1990); Devlin, Science 249: 404-406 (1990); Cwirla et al., Proc. Natl. Acad. Sci. 87: 6378-6382 (1990); Felici, J. Mol. Biol. 222: 301-310 (1991); Ladner *supra*.).

[0113] A compound sometimes alters expression and sometimes alters activity of a polypeptide target and may be a small molecule. Small molecules include, but are not limited to, peptides, peptidomimetics (e.g., peptoids), amino acids, amino acid analogs, polynucleotides, polynucleotide analogs, nucleotides, nucleotide analogs, organic or inorganic compounds (i.e., including heteroorganic and organometallic compounds) having a molecular weight less than about 10,000 grams per mole,

organic or inorganic compounds having a molecular weight less than about 5,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 1,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 500 grams per mole, and salts, esters, and other pharmaceutically acceptable forms of such compounds.

Antisense Nucleic Acid Molecules, Ribozymes, RNAi, siRNA and Modified Nucleic Acid Molecules

[0114] An “antisense” nucleic acid refers to a nucleotide sequence complementary to a “sense” nucleic acid encoding a polypeptide, e.g., complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. The antisense nucleic acid can be complementary to an entire coding strand, or to a portion thereof or a substantially identical sequence thereof. In another embodiment, the antisense nucleic acid molecule is antisense to a “noncoding region” of the coding strand of a nucleotide sequence (e.g., 5' and 3' untranslated regions in SEQ ID NO: 1).

[0115] An antisense nucleic acid can be designed such that it is complementary to the entire coding region of an mRNA encoded by a nucleotide sequence (e.g., SEQ ID NO: 1), and often the antisense nucleic acid is an oligonucleotide antisense to only a portion of a coding or noncoding region of the mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of the mRNA, e.g., between the -10 and +10 regions of the target gene nucleotide sequence of interest. An antisense oligonucleotide can be, for example, about 7, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, or more nucleotides in length. The antisense nucleic acids, which include the ribozymes described hereafter, can be designed to target a *ADAMTS2* nucleotide sequence, often a variant associated with osteoarthritis, or a substantially identical sequence thereof. Among the variants, minor alleles and major alleles can be targeted, and those associated with a higher risk of osteoarthritis are often designed, tested, and administered to subjects.

[0116] An antisense nucleic acid can be constructed using chemical synthesis and enzymatic ligation reactions using standard procedures. For example, an antisense nucleic acid (e.g., an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, e.g., phosphorothioate derivatives and acridine substituted nucleotides can be used. Antisense nucleic acid also can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (i.e., RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

[0117] When utilized as therapeutics, antisense nucleic acids typically are administered to a subject (e.g., by direct injection at a tissue site) or generated in situ such that they hybridize with or bind to

cellular mRNA and/or genomic DNA encoding a polypeptide and thereby inhibit expression of the polypeptide, for example, by inhibiting transcription and/or translation. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then are administered systemically. For systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, for example, by linking antisense nucleic acid molecules to peptides or antibodies which bind to cell surface receptors or antigens. Antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. Sufficient intracellular concentrations of antisense molecules are achieved by incorporating a strong promoter, such as a pol II or pol III promoter, in the vector construct.

[0118] Antisense nucleic acid molecules sometimes are alpha-anomeric nucleic acid molecules. An alpha-anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual beta-units, the strands run parallel to each other (Gaultier et al., Nucleic Acids. Res. 15: 6625-6641 (1987)). Antisense nucleic acid molecules can also comprise a 2'-o-methylribonucleotide (Inoue et al., Nucleic Acids Res. 15: 6131-6148 (1987)) or a chimeric RNA-DNA analogue (Inoue et al., FEBS Lett. 215: 327-330 (1987)). Antisense nucleic acids sometimes are composed of DNA or PNA or any other nucleic acid derivatives described previously.

[0119] In another embodiment, an antisense nucleic acid is a ribozyme. A ribozyme having specificity for a *ADAMTS2* nucleotide sequence can include one or more sequences complementary to such a nucleotide sequence, and a sequence having a known catalytic region responsible for mRNA cleavage (see e.g., U.S. Pat. No. 5,093,246 or Haselhoff and Gerlach, Nature 334: 585-591 (1988)). For example, a derivative of a Tetrahymena L-19 IVS RNA is sometimes utilized in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a mRNA (see e.g., Cech et al. U.S. Patent No. 4,987,071; and Cech et al. U.S. Patent No. 5,116,742). Also, target mRNA sequences can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules (see e.g., Bartel & Szostak, Science 261: 1411-1418 (1993)).

[0120] Osteoarthritis directed molecules include in certain embodiments nucleic acids that can form triple helix structures with a *ADAMTS2* nucleotide sequence, or a substantially identical sequence thereof, especially one that includes a regulatory region that controls expression of a polypeptide. Gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of a nucleotide sequence referenced herein or a substantially identical sequence (e.g., promoter and/or enhancers) to form triple helical structures that prevent transcription of a gene in target cells (see e.g., Helene, Anticancer Drug Des. 6(6): 569-84 (1991); Helene et al., Ann. N.Y. Acad. Sci. 660: 27-36 (1992); and Maher, Bioassays 14(12): 807-15 (1992)). Potential sequences that can be targeted for triple helix formation can be increased by creating a so-called “switchback” nucleic acid molecule. Switchback molecules are synthesized in an alternating 5'-3', 3'-5' manner, such that they base pair with first one

strand of a duplex and then the other, eliminating the necessity for a sizeable stretch of either purines or pyrimidines to be present on one strand of a duplex.

[0121] Osteoarthritis directed molecules include RNAi and siRNA nucleic acids. Gene expression may be inhibited by the introduction of double-stranded RNA (dsRNA), which induces potent and specific gene silencing, a phenomenon called RNA interference or RNAi. See, e.g., Fire et al., US Patent Number 6,506,559; Tuschl et al. PCT International Publication No. WO 01/75164; Kay et al. PCT International Publication No. WO 03/010180A1; or Bosher JM, Labouesse, Nat Cell Biol 2000 Feb;2(2):E31-6. This process has been improved by decreasing the size of the double-stranded RNA to 20-24 base pairs (to create small-interfering RNAs or siRNAs) that “switched off” genes in mammalian cells without initiating an acute phase response, i.e., a host defense mechanism that often results in cell death (see, e.g., Caplen et al. Proc Natl Acad Sci U S A. 2001 Aug 14;98(17):9742-7 and Elbashir et al. Methods 2002 Feb;26(2):199-213). There is increasing evidence of post-transcriptional gene silencing by RNA interference (RNAi) for inhibiting targeted expression in mammalian cells at the mRNA level, in human cells. There is additional evidence of effective methods for inhibiting the proliferation and migration of tumor cells in human patients, and for inhibiting metastatic cancer development (see, e.g., U.S. Patent Application No. US2001000993183; Caplen et al. Proc Natl Acad Sci U S A; and Abderrahmani et al. Mol Cell Biol 2001 Nov21(21):7256-67).

[0122] An “siRNA” or “RNAi” refers to a nucleic acid that forms a double stranded RNA and has the ability to reduce or inhibit expression of a gene or target gene when the siRNA is delivered to or expressed in the same cell as the gene or target gene. “siRNA” refers to short double-stranded RNA formed by the complementary strands. Complementary portions of the siRNA that hybridize to form the double stranded molecule often have substantial or complete identity to the target molecule sequence. In one embodiment, an siRNA refers to a nucleic acid that has substantial or complete identity to a target gene and forms a double stranded siRNA.

[0123] When designing the siRNA molecules, the targeted region often is selected from a given DNA sequence beginning 50 to 100 nucleotides downstream of the start codon. See, e.g., Elbashir et al., Methods 26:199-213 (2002). Initially, 5' or 3' UTRs and regions nearby the start codon were avoided assuming that UTR-binding proteins and/or translation initiation complexes may interfere with binding of the siRNP or RISC endonuclease complex. Sometimes regions of the target 23 nucleotides in length conforming to the sequence motif AA(N19)TT (N, an nucleotide), and regions with approximately 30% to 70% G/C-content (often about 50% G/C-content) often are selected. If no suitable sequences are found, the search often is extended using the motif NA(N21). The sequence of the sense siRNA sometimes corresponds to (N19) TT or N21 (position 3 to 23 of the 23-nt motif), respectively. In the latter case, the 3' end of the sense siRNA often is converted to TT. The rationale for this sequence conversion is to generate a symmetric duplex with respect to the sequence composition of the sense and

antisense 3' overhangs. The antisense siRNA is synthesized as the complement to position 1 to 21 of the 23-nt motif. Because position 1 of the 23-nt motif is not recognized sequence-specifically by the antisense siRNA, the 3'-most nucleotide residue of the antisense siRNA can be chosen deliberately. However, the penultimate nucleotide of the antisense siRNA (complementary to position 2 of the 23-nt motif) often is complementary to the targeted sequence. For simplifying chemical synthesis, TT often is utilized. siRNAs corresponding to the target motif NAR(N17)YNN, where R is purine (A,G) and Y is pyrimidine (C,U), often are selected. Respective 21 nucleotide sense and antisense siRNAs often begin with a purine nucleotide and can also be expressed from pol III expression vectors without a change in targeting site. Expression of RNAs from pol III promoters often is efficient when the first transcribed nucleotide is a purine.

[0124] The sequence of the siRNA can correspond to the full length target gene, or a subsequence thereof. Often, the siRNA is about 15 to about 50 nucleotides in length (e.g., each complementary sequence of the double stranded siRNA is 15-50 nucleotides in length, and the double stranded siRNA is about 15-50 base pairs in length, sometimes about 20-30 nucleotides in length or about 20-25 nucleotides in length, e.g., 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 nucleotides in length. The siRNA sometimes is about 21 nucleotides in length. Methods of using siRNA are well known in the art, and specific siRNA molecules may be purchased from a number of companies including Dharmacon Research, Inc.

[0125] Antisense, ribozyme, RNAi and siRNA nucleic acids can be altered to form modified nucleic acid molecules. The nucleic acids can be altered at base moieties, sugar moieties or phosphate backbone moieties to improve stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of nucleic acid molecules can be modified to generate peptide nucleic acids (see Hyrup et al., Bioorganic & Medicinal Chemistry 4 (1): 5-23 (1996)). As used herein, the terms “peptide nucleic acid” or “PNA” refers to a nucleic acid mimic such as a DNA mimic, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of a PNA can allow for specific hybridization to DNA and RNA under conditions of low ionic strength. Synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described, for example, in Hyrup et al., (1996) *supra* and Perry-O’Keefe et al., Proc. Natl. Acad. Sci. 93: 14670-675 (1996).

[0126] PNA nucleic acids can be used in prognostic, diagnostic, and therapeutic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, for example, inducing transcription or translation arrest or inhibiting replication. PNA nucleic acid molecules can also be used in the analysis of single base pair mutations in a gene, (e.g., by PNA-directed PCR clamping); as “artificial restriction enzymes” when used in combination with other enzymes, (e.g., S1 nucleases (Hyrup (1996) *supra*)); or as probes or primers for DNA sequencing or hybridization (Hyrup et al., (1996) *supra*; Perry-O’Keefe *supra*).

[0127] In other embodiments, oligonucleotides may include other appended groups such as peptides (e.g., for targeting host cell receptors *in vivo*), or agents facilitating transport across cell membranes (see e.g., Letsinger et al., Proc. Natl. Acad. Sci. USA 86: 6553-6556 (1989); Lemaitre et al., Proc. Natl. Acad. Sci. USA 84: 648-652 (1987); PCT Publication No. W088/09810) or the blood-brain barrier (see, e.g., PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization-triggered cleavage agents (See, e.g., Krol et al., Bio-Techniques 6: 958-976 (1988)) or intercalating agents. (See, e.g., Zon, Pharm. Res. 5: 539-549 (1988)). To this end, the oligonucleotide may be conjugated to another molecule, (e.g., a peptide, hybridization triggered cross-linking agent, transport agent, or hybridization-triggered cleavage agent).

[0128] Also included herein are molecular beacon oligonucleotide primer and probe molecules having one or more regions complementary to a *ADAMTS2* nucleotide sequence, or a substantially identical sequence thereof, two complementary regions one having a fluorophore and one a quencher such that the molecular beacon is useful for quantifying the presence of the nucleic acid in a sample. Molecular beacon nucleic acids are described, for example, in Lizardi et al., U.S. Patent No. 5,854,033; Nazarenko et al., U.S. Patent No. 5,866,336, and Livak et al., U.S. Patent 5,876,930.

Antibodies

[0129] The term “antibody” as used herein refers to an immunoglobulin molecule or immunologically active portion thereof, i.e., an antigen-binding portion. Examples of immunologically active portions of immunoglobulin molecules include F(ab) and F(ab')₂ fragments which can be generated by treating the antibody with an enzyme such as pepsin. An antibody sometimes is a polyclonal, monoclonal, recombinant (e.g., a chimeric or humanized), fully human, non-human (e.g., murine), or a single chain antibody. An antibody may have effector function and can fix complement, and is sometimes coupled to a toxin or imaging agent.

[0130] A full-length polypeptide or antigenic peptide fragment encoded by a nucleotide sequence referenced herein can be used as an immunogen or can be used to identify antibodies made with other immunogens, e.g., cells, membrane preparations, and the like. An antigenic peptide often includes at least 8 amino acid residues of the amino acid sequences encoded by a nucleotide sequence referenced herein, or substantially identical sequence thereof, and encompasses an epitope. Antigenic peptides sometimes include 10 or more amino acids, 15 or more amino acids, 20 or more amino acids, or 30 or more amino acids. Hydrophilic and hydrophobic fragments of polypeptides sometimes are used as immunogens.

[0131] Epitopes encompassed by the antigenic peptide are regions located on the surface of the polypeptide (e.g., hydrophilic regions) as well as regions with high antigenicity. For example, an Emini surface probability analysis of the human polypeptide sequence can be used to indicate the regions that

have a particularly high probability of being localized to the surface of the polypeptide and are thus likely to constitute surface residues useful for targeting antibody production. The antibody may bind an epitope on any domain or region on polypeptides described herein.

[0132] Also, chimeric, humanized, and completely human antibodies are useful for applications which include repeated administration to subjects. Chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, can be made using standard recombinant DNA techniques. Such chimeric and humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in Robinson et al International Application No. PCT/US86/02269; Akira, et al European Patent Application 184,187; Taniguchi, M., European Patent Application 171,496; Morrison et al European Patent Application 173,494; Neuberger et al PCT International Publication No. WO 86/01533; Cabilly et al U.S. Patent No. 4,816,567; Cabilly et al European Patent Application 125,023; Better et al., Science 240: 1041-1043 (1988); Liu et al., Proc. Natl. Acad. Sci. USA 84: 3439-3443 (1987); Liu et al., J. Immunol. 139: 3521-3526 (1987); Sun et al., Proc. Natl. Acad. Sci. USA 84: 214-218 (1987); Nishimura et al., Canc. Res. 47: 999-1005 (1987); Wood et al., Nature 314: 446-449 (1985); and Shaw et al., J. Natl. Cancer Inst. 80: 1553-1559 (1988); Morrison, S. L., Science 229: 1202-1207 (1985); Oi et al., BioTechniques 4: 214 (1986); Winter U.S. Patent 5,225,539; Jones et al., Nature 321: 552-525 (1986); Verhoeven et al., Science 239: 1534; and Beidler et al., J. Immunol. 141: 4053-4060 (1988).

[0133] Completely human antibodies are particularly desirable for therapeutic treatment of human patients. Such antibodies can be produced using transgenic mice that are incapable of expressing endogenous immunoglobulin heavy and light chains genes, but which can express human heavy and light chain genes. See, for example, Lonberg and Huszar, Int. Rev. Immunol. 13: 65-93 (1995); and U.S. Patent Nos. 5,625,126; 5,633,425; 5,569,825; 5,661,016; and 5,545,806. In addition, companies such as Abgenix, Inc. (Fremont, CA) and Medarex, Inc. (Princeton, NJ), can be engaged to provide human antibodies directed against a selected antigen using technology similar to that described above. Completely human antibodies that recognize a selected epitope also can be generated using a technique referred to as "guided selection." In this approach a selected non-human monoclonal antibody (e.g., a murine antibody) is used to guide the selection of a completely human antibody recognizing the same epitope. This technology is described for example by Jespers et al., Bio/Technology 12: 899-903 (1994).

[0134] An antibody can be a single chain antibody. A single chain antibody (scFV) can be engineered (see, e.g., Colcher et al., Ann. N Y Acad. Sci. 880: 263-80 (1999); and Reiter, Clin. Cancer Res. 2: 245-52 (1996)). Single chain antibodies can be dimerized or multimerized to generate multivalent antibodies having specificities for different epitopes of the same target polypeptide.

[0135] Antibodies also may be selected or modified so that they exhibit reduced or no ability to bind an Fc receptor. For example, an antibody may be an isotype or subtype, fragment or other mutant, which

does not support binding to an Fc receptor (e.g., it has a mutagenized or deleted Fc receptor binding region).

[0136] Also, an antibody (or fragment thereof) may be conjugated to a therapeutic moiety such as a cytotoxin, a therapeutic agent or a radioactive metal ion. A cytotoxin or cytotoxic agent includes any agent that is detrimental to cells. Examples include taxol, cytochalasin B, gramicidin D, ethidium bromide, emetine, mitomycin, etoposide, tenoposide, vincristine, vinblastine, colchicin, doxorubicin, daunorubicin, dihydroxy anthracin dione, mitoxantrone, mithramycin, actinomycin D, 1 dehydrotestosterone, glucocorticoids, procaine, tetracaine, lidocaine, propranolol, and puromycin and analogs or homologs thereof. Therapeutic agents include, but are not limited to, antimetabolites (e.g., methotrexate, 6-mercaptopurine, 6-thioguanine, cytarabine, 5-fluorouracil decarbazine), alkylating agents (e.g., mechlorethamine, thiotepa chlorambucil, melphalan, carmustine (BCNU) and lomustine (CCNU), cyclophosphamide, busulfan, dibromomannitol, streptozotocin, mitomycin C, and cis-dichlorodiamine platinum (II) (DDP) cisplatin), anthracyclines (e.g., daunorubicin (formerly daunomycin) and doxorubicin), antibiotics (e.g., dactinomycin (formerly actinomycin), bleomycin, mithramycin, and anthramycin (AMC)), and anti-mitotic agents (e.g., vincristine and vinblastine).

[0137] Antibody conjugates can be used for modifying a given biological response. For example, the drug moiety may be a protein or polypeptide possessing a desired biological activity. Such proteins may include, for example, a toxin such as abrin, ricin A, pseudomonas exotoxin, or diphtheria toxin; a polypeptide such as tumor necrosis factor, gamma-interferon, alpha-interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator; or, biological response modifiers such as, for example, lymphokines, interleukin-1 ("IL-1"), interleukin-2 ("IL-2"), interleukin-6 ("IL-6"), granulocyte macrophage colony stimulating factor ("GM-CSF"), granulocyte colony stimulating factor ("G-CSF"), or other growth factors. Also, an antibody can be conjugated to a second antibody to form an antibody heteroconjugate as described by Segal in U.S. Patent No. 4,676,980, for example.

[0138] An antibody (e.g., monoclonal antibody) can be used to isolate target polypeptides by standard techniques, such as affinity chromatography or immunoprecipitation. Moreover, an antibody can be used to detect a target polypeptide (e.g., in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the polypeptide. Antibodies can be used diagnostically to monitor polypeptide levels in tissue as part of a clinical testing procedure, e.g., to determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (i.e., physically linking) the antibody to a detectable substance (i.e., antibody labeling). Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β -galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials

include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ¹²⁵I, ¹³¹I, ³⁵S or ³H. Also, an antibody can be utilized as a test molecule for determining whether it can treat osteoarthritis, and as a therapeutic for administration to a subject for treating osteoarthritis.

[0139] An antibody can be made by immunizing with a purified antigen, or a fragment thereof, e.g., a fragment described herein, a membrane associated antigen, tissues, e.g., crude tissue preparations, whole cells, preferably living cells, lysed cells, or cell fractions.

[0140] Included herein are antibodies which bind only a native polypeptide, only denatured or otherwise non-native polypeptide, or which bind both, as well as those having linear or conformational epitopes. Conformational epitopes sometimes can be identified by selecting antibodies that bind to native but not denatured polypeptide. Also featured are antibodies that specifically bind to a polypeptide variant associated with osteoarthritis.

Methods for Identifying Candidate Therapeutics for Treating Osteoarthritis

[0141] Current therapies for the treatment of osteoarthritis have limited efficacy, limited tolerability and significant mechanism-based side effects, and few of the available therapies adequately address underlying defects. Current therapeutic approaches were largely developed in the absence of defined molecular targets or even a solid understanding of disease pathogenesis. Therefore, provided are methods of identifying candidate therapeutics that target biochemical pathways related to the development of osteoarthritis.

[0142] Thus, featured herein are methods for identifying a candidate therapeutic for treating osteoarthritis. The methods comprise contacting a test molecule with a target molecule in a system. A "target molecule" as used herein refers to a *ADAMTS2* nucleic acid, a substantially identical nucleic acid thereof, or a fragment thereof, and an encoded polypeptide of the foregoing. The methods also comprise determining the presence or absence of an interaction between the test molecule and the target molecule, where the presence of an interaction between the test molecule and the nucleic acid or polypeptide identifies the test molecule as a candidate osteoarthritis therapeutic. The interaction between the test molecule and the target molecule may be quantified.

[0143] Test molecules and candidate therapeutics include, but are not limited to, compounds, antisense nucleic acids, siRNA molecules, ribozymes, polypeptides or proteins encoded by a *ADAMTS2* nucleotide sequence, or a substantially identical sequence or fragment thereof, and immunotherapeutics (e.g., antibodies and HLA-presented polypeptide fragments). A test molecule or candidate therapeutic may act as a modulator of target molecule concentration or target molecule function in a system. A

“modulator” may agonize (i.e., up-regulates) or antagonize (i.e., down-regulates) a target molecule concentration partially or completely in a system by affecting such cellular functions as DNA replication and/or DNA processing (e.g., DNA methylation or DNA repair), RNA transcription and/or RNA processing (e.g., removal of intronic sequences and/or translocation of spliced mRNA from the nucleus), polypeptide production (e.g., translation of the polypeptide from mRNA), and/or polypeptide post-translational modification (e.g., glycosylation, phosphorylation, and proteolysis of pro-polypeptides). A modulator may also agonize or antagonize a biological function of a target molecule partially or completely, where the function may include adopting a certain structural conformation, interacting with one or more binding partners, ligand binding, catalysis (e.g., phosphorylation, dephosphorylation, hydrolysis, methylation, and isomerization), and an effect upon a cellular event (e.g., effecting progression of osteoarthritis).

[0144] As used herein, the term “system” refers to a cell free *in vitro* environment and a cell-based environment such as a collection of cells, a tissue, an organ, or an organism. A system is “contacted” with a test molecule in a variety of manners, including adding molecules in solution and allowing them to interact with one another by diffusion, cell injection, and any administration routes in an animal. As used herein, the term “interaction” refers to an effect of a test molecule on test molecule, where the effect sometimes is binding between the test molecule and the target molecule, and sometimes is an observable change in cells, tissue, or organism.

[0145] There are many standard methods for detecting the presence or absence of interaction between a test molecule and a target molecule. For example, titrometric, acidimetric, radiometric, NMR, monolayer, polarographic, spectrophotometric, fluorescent, and ESR assays probative of a target molecule interaction may be utilized.

[0146] *ADAMTS2* activity and/or *ADAMTS2* interactions can be detected and quantified using assays known in the art. For example, an immunoprecipitation assay or a kinase activity assay that employs a kinase-inactivated MEK can be utilized. Kinase inactivated MEKs are known in the art, such as a MEK that includes the mutation K97M. In these assays, mammalian cells (e.g., COS or NIH-3T3) are transiently transfected with constructs expressing *ADAMTS2*, and in addition, the cells are co-transfected with oncogenic RAS or SRC or both. Oncogenic RAS or SRC activates ADAMTS2 kinase activity. *ADAMTS2* is immunoprecipitated from cell extracts using a monoclonal antibody (e.g., 9E10) or a polyclonal antibody (e.g., from rabbit) specific for a unique peptide from *ADAMTS2*. *ADAMTS2* is then resuspended in assay buffer containing GST-Mek1 or GST-Mek2 and/or GST-ERK2. In addition, [γ ³²P] ATP can be added to detect and/or quantify phosphorylation activity. Samples are incubated for 5-30 minutes at 30°C, and then the reaction is terminated by addition of EDTA. The samples are centrifuged and the supernatant fractions are collected. Phosphorylation activity is detected using one of two methods: (i) activity of GST-ERK2 kinase can be measured using MBP (myelin basic

protein, a substrate for ERK) as substrate, or (ii) following incubation of immunoprecipitated *ADAMTS2* in reaction buffer containing GST-ERK and [gamma ³²P] ATP, transfer of labeled ATP to kinase-dead ERK can be quantified by a phosphor-imager or densitometer following PAGE separation of polypeptide products (phosphorylated and non-phosphorylated forms). These types of assays are described in Weber et al., *Oncogene* 19: 169-176 (2000); Mason et al., *EMBO J.* 18: 2137-2148 (1999); Marais et al., *J. Biol. Chem.* 272: 4378-4383 (1997); Marais et al., *EMBO J.* 14: 3136-3145 (1995).

[0147] As noted above, *ADAMTS2* includes a domain having metalloprotease activity, and modulators of such activity are known. Examples of such modulators are set forth in WO03063762A2; WO-09937625; WO-09918076; WO-09838163; WO-09837877; WO9947550A1; WO0177092A1; WO0040577A1; WO9942436A1; WO9838163A1; WO9837877A1; WO04014379A1; WO03106381A2; WO03014098A1; WO03014092A1 and WO02096426A1.

[0148] Test molecule/target molecule interactions can be detected and/or quantified using assays known in the art. For example, an interaction can be determined by labeling the test molecule and/or the target molecule, where the label is covalently or non-covalently attached to the test molecule or target molecule. The label is sometimes a radioactive molecule such as ¹²⁵I, ¹³¹I, ³⁵S or ³H, which can be detected by direct counting of radioemission or by scintillation counting. Also, enzymatic labels such as horseradish peroxidase, alkaline phosphatase, or luciferase may be utilized where the enzymatic label can be detected by determining conversion of an appropriate substrate to product. In addition, presence or absence of an interaction can be determined without labeling. For example, a microphysiometer (*e.g.*, Cytosensor) is an analytical instrument that measures the rate at which a cell acidifies its environment using a light-addressable potentiometric sensor (LAPS). Changes in this acidification rate can be used as an indication of an interaction between a test molecule and target molecule (McConnell, H. M. et al., *Science* 257: 1906-1912 (1992)).

[0149] In cell-based systems, cells typically include a *ADAMTS2* nucleic acid, an encoded polypeptide, or substantially identical nucleic acid or polypeptide thereof, and are often of mammalian origin, although the cell can be of any origin. Whole cells, cell homogenates, and cell fractions (*e.g.*, cell membrane fractions) can be subjected to analysis. Where interactions between a test molecule with a target polypeptide are monitored, soluble and/or membrane bound forms of the polypeptide may be utilized. Where membrane-bound forms of the polypeptide are used, it may be desirable to utilize a solubilizing agent. Examples of such solubilizing agents include non-ionic detergents such as n-octylglucoside, n-dodecylglucoside, n-dodecylmaltoside, octanoyl-N-methylglucamide, decanoyl-N-methylglucamide, Triton® X-100, Triton® X-114, Thesit®, Isotridecyloxy(ethylene glycol ether)_n, 3-[(3-cholamidopropyl)dimethylammonio]-1-propane sulfonate (CHAPS), 3-[(3-

cholamidopropyl)dimethylamminio]-2-hydroxy-1-propane sulfonate (CHAPSO), or N-dodecyl-N,N-dimethyl-3-ammonio-1-propane sulfonate.

[0150] An interaction between a test molecule and target molecule also can be detected by monitoring fluorescence energy transfer (FET) (*see, e.g.*, Lakowicz *et al.*, U.S. Patent No. 5,631,169; Stavrianopoulos *et al.* U.S. Patent No. 4,868,103). A fluorophore label on a first, “donor” molecule is selected such that its emitted fluorescent energy will be absorbed by a fluorescent label on a second, “acceptor” molecule, which in turn is able to fluoresce due to the absorbed energy. Alternately, the “donor” polypeptide molecule may simply utilize the natural fluorescent energy of tryptophan residues. Labels are chosen that emit different wavelengths of light, such that the “acceptor” molecule label may be differentiated from that of the “donor”. Since the efficiency of energy transfer between the labels is related to the distance separating the molecules, the spatial relationship between the molecules can be assessed. In a situation in which binding occurs between the molecules, the fluorescent emission of the “acceptor” molecule label in the assay should be maximal. An FET binding event can be conveniently measured through standard fluorometric detection means well known in the art (*e.g.*, using a fluorimeter).

[0151] In another embodiment, determining the presence or absence of an interaction between a test molecule and a target molecule can be effected by monitoring surface plasmon resonance (*see, e.g.*, Sjolander & Urbaniczky, *Anal. Chem.* 63: 2338-2345 (1991) and Szabo *et al.*, *Curr. Opin. Struct. Biol.* 5: 699-705 (1995)). “Surface plasmon resonance” or “biomolecular interaction analysis (BIA)” can be utilized to detect biospecific interactions in real time, without labeling any of the interactants (*e.g.*, BIACore). Changes in the mass at the binding surface (indicative of a binding event) result in alterations of the refractive index of light near the surface (the optical phenomenon of surface plasmon resonance (SPR)), resulting in a detectable signal which can be used as an indication of real-time reactions between biological molecules.

[0152] In another embodiment, the target molecule or test molecules are anchored to a solid phase, facilitating the detection of target molecule/test molecule complexes and separation of the complexes from free, uncomplexed molecules. The target molecule or test molecule is immobilized to the solid support. In an embodiment, the target molecule is anchored to a solid surface, and the test molecule, which is not anchored, can be labeled, either directly or indirectly, with detectable labels discussed herein.

[0153] It may be desirable to immobilize a target molecule, an anti-target molecule antibody, and/or test molecules to facilitate separation of target molecule/test molecule complexes from uncomplexed forms, as well as to accommodate automation of the assay. The attachment between a test molecule and/or target molecule and the solid support may be covalent or non-covalent (*see, e.g.*, U.S. Patent No. 6,022,688 for non-covalent attachments). The solid support may be one or more surfaces of the system, such as one or more surfaces in each well of a microtiter plate, a surface of a silicon wafer, a surface of a

bead (*see, e.g.*, Lam, *Nature* 354: 82-84 (1991)) that is optionally linked to another solid support, or a channel in a microfluidic device, for example. Types of solid supports, linker molecules for covalent and non-covalent attachments to solid supports, and methods for immobilizing nucleic acids and other molecules to solid supports are well known (*see, e.g.*, U.S. Patent Nos. 6,261,776; 5,900,481; 6,133,436; and 6,022,688; and WIPO publication WO 01/18234).

[0154] In an embodiment, target molecule may be immobilized to surfaces via biotin and streptavidin. For example, biotinylated target polypeptide can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques known in the art (*e.g.*, biotinylation kit, Pierce Chemicals, Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). In another embodiment, a target polypeptide can be prepared as a fusion polypeptide. For example, glutathione-S-transferase/target polypeptide fusion can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivitized microtiter plates, which are then combined with a test molecule under conditions conducive to complex formation (*e.g.*, at physiological conditions for salt and pH). Following incubation, the beads or microtiter plate wells are washed to remove any unbound components, or the matrix is immobilized in the case of beads, and complex formation is determined directly or indirectly as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of target molecule binding or activity is determined using standard techniques.

[0155] In an embodiment, the non-immobilized component is added to the coated surface containing the anchored component. After the reaction is complete, unreacted components are removed (*e.g.*, by washing) under conditions such that a significant percentage of complexes formed will remain immobilized to the solid surface. The detection of complexes anchored on the solid surface can be accomplished in a number of manners. Where the previously non-immobilized component is pre-labeled, the detection of label immobilized on the surface indicates that complexes were formed. Where the previously non-immobilized component is not pre-labeled, an indirect label can be used to detect complexes anchored on the surface, *e.g.*, by adding a labeled antibody specific for the immobilized component, where the antibody, in turn, can be directly labeled or indirectly labeled with, *e.g.*, a labeled anti-Ig antibody.

[0156] In another embodiment, an assay is performed utilizing antibodies that specifically bind target molecule or test molecule but do not interfere with binding of the target molecule to the test molecule. Such antibodies can be derivitized to a solid support, and unbound target molecule may be immobilized by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the target molecule, as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the target molecule.

[0157] Cell free assays also can be conducted in a liquid phase. In such an assay, reaction products are separated from unreacted components, by any of a number of standard techniques, including but not limited to: differential centrifugation (*see, e.g.*, Rivas, G., and Minton, *Trends Biochem Sci Aug;18(8): 284-7 (1993)*); chromatography (gel filtration chromatography, ion-exchange chromatography); electrophoresis (*see, e.g.*, Ausubel *et al.*, eds. *Current Protocols in Molecular Biology*, J. Wiley: New York (1999)); and immunoprecipitation (*see, e.g.*, Ausubel *et al.*, eds., *supra*). Media and chromatographic techniques are known to one skilled in the art (*see, e.g.*, Heegaard, *J Mol. Recognit. Winter; 11(1-6): 141-8 (1998)*; Hage & Tweed, *J. Chromatogr. B Biomed. Sci. Appl. Oct 10; 699 (1-2): 499-525 (1997)*). Further, fluorescence energy transfer may also be conveniently utilized, as described herein, to detect binding without further purification of the complex from solution.

[0158] In another embodiment, modulators of target molecule expression are identified. For example, a cell or cell free mixture is contacted with a candidate compound and the expression of target mRNA or target polypeptide is evaluated relative to the level of expression of target mRNA or target polypeptide in the absence of the candidate compound. When expression of target mRNA or target polypeptide is greater in the presence of the candidate compound than in its absence, the candidate compound is identified as an agonist of target mRNA or target polypeptide expression. Alternatively, when expression of target mRNA or target polypeptide is less (*e.g.*, less with statistical significance) in the presence of the candidate compound than in its absence, the candidate compound is identified as an antagonist or inhibitor of target mRNA or target polypeptide expression. The level of target mRNA or target polypeptide expression can be determined by methods described herein.

[0159] In another embodiment, binding partners that interact with a target molecule are detected. The target molecules can interact with one or more cellular or extracellular macromolecules, such as polypeptides *in vivo*, and these interacting molecules are referred to herein as "binding partners." Binding partners can agonize or antagonize target molecule biological activity. Also, test molecules that agonize or antagonize interactions between target molecules and binding partners can be useful as therapeutic molecules as they can up-regulate or down-regulated target molecule activity *in vivo* and thereby treat osteoarthritis.

[0160] Binding partners of target molecules can be identified by methods known in the art. For example, binding partners may be identified by lysing cells and analyzing cell lysates by electrophoretic techniques. Alternatively, a two-hybrid assay or three-hybrid assay can be utilized (*see, e.g.*, U.S. Patent No. 5,283,317; Zervos *et al.*, *Cell 72:223-232 (1993)*; Madura *et al.*, *J. Biol. Chem. 268: 12046-12054 (1993)*; Bartel *et al.*, *Biotechniques 14: 920-924 (1993)*; Iwabuchi *et al.*, *Oncogene 8: 1693-1696 (1993)*; and Brent WO94/10300). A two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. The assay often utilizes two different DNA constructs. In one construct, a *ADAMTS2* nucleic acid (sometimes referred to as the

“bait”) is fused to a gene encoding the DNA binding domain of a known transcription factor (*e.g.*, GAL-4). In another construct, a DNA sequence from a library of DNA sequences that encodes a potential binding partner (sometimes referred to as the “prey”) is fused to a gene that encodes an activation domain of the known transcription factor. Sometimes, a *ADAMTS2* nucleic acid can be fused to the activation domain. If the “bait” and the “prey” molecules interact *in vivo*, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (*e.g.*, LacZ) which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to identify the potential binding partner.

[0161] In an embodiment for identifying test molecules that antagonize or agonize complex formation between target molecules and binding partners, a reaction mixture containing the target molecule and the binding partner is prepared, under conditions and for a time sufficient to allow complex formation. The reaction mixture often is provided in the presence or absence of the test molecule. The test molecule can be included initially in the reaction mixture, or can be added at a time subsequent to the addition of the target molecule and its binding partner. Control reaction mixtures are incubated without the test molecule or with a placebo. Formation of any complexes between the target molecule and the binding partner then is detected. Decreased formation of a complex in the reaction mixture containing test molecule as compared to in a control reaction mixture indicates that the molecule antagonizes target molecule/binding partner complex formation. Alternatively, increased formation of a complex in the reaction mixture containing test molecule as compared to in a control reaction mixture indicates that the molecule agonizes target molecule/binding partner complex formation. In another embodiment, complex formation of target molecule/binding partner can be compared to complex formation of mutant target molecule/binding partner (*e.g.*, amino acid modifications in a target polypeptide). Such a comparison can be important in those cases where it is desirable to identify test molecules that modulate interactions of mutant but not non-mutated target gene products.

[0162] The assays can be conducted in a heterogeneous or homogeneous format. In heterogeneous assays, target molecule and/or the binding partner are immobilized to a solid phase, and complexes are detected on the solid phase at the end of the reaction. In homogeneous assays, the entire reaction is carried out in a liquid phase. In either approach, the order of addition of reactants can be varied to obtain different information about the molecules being tested. For example, test compounds that agonize target molecule/binding partner interactions can be identified by conducting the reaction in the presence of the test molecule in a competition format. Alternatively, test molecules that agonize preformed complexes, *e.g.*, molecules with higher binding constants that displace one of the components from the complex, can be tested by adding the test compound to the reaction mixture after complexes have been formed.

[0163] In a heterogeneous assay embodiment, the target molecule or the binding partner is anchored onto a solid surface (e.g., a microtiter plate), while the non-anchored species is labeled, either directly or indirectly. The anchored molecule can be immobilized by non-covalent or covalent attachments. Alternatively, an immobilized antibody specific for the molecule to be anchored can be used to anchor the molecule to the solid surface. The partner of the immobilized species is exposed to the coated surface with or without the test molecule. After the reaction is complete, unreacted components are removed (e.g., by washing) such that a significant portion of any complexes formed will remain immobilized on the solid surface. Where the non-immobilized species is pre-labeled, the detection of label immobilized on the surface is indicative of complex. Where the non-immobilized species is not pre-labeled, an indirect label can be used to detect complexes anchored to the surface; e.g., by using a labeled antibody specific for the initially non-immobilized species. Depending upon the order of addition of reaction components, test compounds that inhibit complex formation or that disrupt preformed complexes can be detected.

[0164] In another embodiment, the reaction can be conducted in a liquid phase in the presence or absence of test molecule, where the reaction products are separated from unreacted components, and the complexes are detected (e.g., using an immobilized antibody specific for one of the binding components to anchor any complexes formed in solution, and a labeled antibody specific for the other partner to detect anchored complexes). Again, depending upon the order of addition of reactants to the liquid phase, test compounds that inhibit complex or that disrupt preformed complexes can be identified.

[0165] In an alternate embodiment, a homogeneous assay can be utilized. For example, a preformed complex of the target gene product and the interactive cellular or extracellular binding partner product is prepared. One or both of the target molecule or binding partner is labeled, and the signal generated by the label(s) is quenched upon complex formation (e.g., U.S. Patent No. 4,109,496 that utilizes this approach for immunoassays). Addition of a test molecule that competes with and displaces one of the species from the preformed complex will result in the generation of a signal above background. In this way, test substances that disrupt target molecule/binding partner complexes can be identified.

[0166] Candidate therapeutics for treating osteoarthritis are identified from a group of test molecules that interact with a target molecule. Test molecules are normally ranked according to the degree with which they modulate (e.g., agonize or antagonize) a function associated with the target molecule (e.g., DNA replication and/or processing, RNA transcription and/or processing, polypeptide production and/or processing, and/or biological function/activity), and then top ranking modulators are selected. Also, pharmacogenomic information described herein can determine the rank of a modulator. The top 10% of ranked test molecules often are selected for further testing as candidate therapeutics, and sometimes the top 15%, 20%, or 25% of ranked test molecules are selected for further testing as candidate therapeutics. Candidate therapeutics typically are formulated for administration to a subject.

Therapeutic Formulations

[0167] Formulations and pharmaceutical compositions typically include in combination with a pharmaceutically acceptable carrier one or more target molecule modulators. The modulator often is a test molecule identified as having an interaction with a target molecule by a screening method described above. The modulator may be a compound, an antisense nucleic acid, a ribozyme, an antibody, or a binding partner. Also, formulations may comprise a target polypeptide or fragment thereof in combination with a pharmaceutically acceptable carrier.

[0168] Formulations or pharmaceutical compositions typically include in combination with a pharmaceutically acceptable carrier, a compound, an antisense nucleic acid, a ribozyme, an antibody, a binding partner that interacts with an *ADAMTS2* polypeptide, a *ADAMTS2* nucleic acid, or a fragment thereof. The formulated molecule may be one that is identified by a screening method described above. Also, formulations may comprise a *ADAMTS2* polypeptide or fragment thereof, where the *ADAMTS2* polypeptide contains an isoleucine at position 245 of SEQ ID NO: 4, and a pharmaceutically acceptable carrier. Also, formulations may comprise an active *ADAMTS2* polypeptide or fragment thereof, where *ADAMTS2* polypeptide fragments having activity are selected from amino acids 252-1211, 253-1211, 254-1211, 255-1211, 256-1211, 257-1211, 258-1211, 259-1211 or 260-1211 of SEQ ID NO: 4, where it is understood that the active form of *ADAMTS2* does not contain the propeptide domain. As used herein, the term "pharmaceutically acceptable carrier" includes solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. Supplementary active compounds can also be incorporated into the compositions.

[0169] As used herein, the term "pharmaceutically acceptable carrier" includes solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. Supplementary active compounds can also be incorporated into the compositions. Pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

[0170] A pharmaceutical composition typically is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, e.g., intravenous, intradermal, subcutaneous, oral (e.g., inhalation), transdermal (topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerin, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates and agents for the adjustment of tonicity such as sodium chloride or dextrose. pH can be adjusted with

acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

[0171] Oral compositions generally include an inert diluent or an edible carrier. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules, e.g., gelatin capsules. Oral compositions can also be prepared using a fluid carrier for use as a mouthwash. Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

[0172] Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™ (BASF, Parsippany, NJ) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringability exists. It should be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as mannitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

[0173] Sterile injectable solutions can be prepared by incorporating the active compound in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle which contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum drying and freeze-drying which

yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

[0174] For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable propellant, *e.g.*, a gas such as carbon dioxide, or a nebulizer.

[0175] Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art. Molecules can also be prepared in the form of suppositories (*e.g.*, with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

[0176] In one embodiment, active molecules are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. Materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Patent No. 4,522,811.

[0177] It is advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier.

[0178] Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD₅₀ (the dose lethal to 50% of the population) and the ED₅₀ (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD₅₀/ED₅₀. Molecules which exhibit high therapeutic indices are preferred. While molecules that exhibit toxic side effects may be used, care should be taken to design a delivery system

that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

[0179] The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such molecules lies preferably within a range of circulating concentrations that include the ED₅₀ with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any molecules used in the methods described herein, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC₅₀ (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

[0180] As defined herein, a therapeutically effective amount of protein or polypeptide (*i.e.*, an effective dosage) ranges from about 0.001 to 30 mg/kg body weight, sometimes about 0.01 to 25 mg/kg body weight, often about 0.1 to 20 mg/kg body weight, and more often about 1 to 10 mg/kg, 2 to 9 mg/kg, 3 to 8 mg/kg, 4 to 7 mg/kg, or 5 to 6 mg/kg body weight. The protein or polypeptide can be administered one time per week for between about 1 to 10 weeks, sometimes between 2 to 8 weeks, often between about 3 to 7 weeks, and more often for about 4, 5, or 6 weeks. The skilled artisan will appreciate that certain factors may influence the dosage and timing required to effectively treat a subject, including but not limited to the severity of the disease or disorder, previous treatments, the general health and/or age of the subject, and other diseases present. Moreover, treatment of a subject with a therapeutically effective amount of a protein, polypeptide, or antibody can include a single treatment or, preferably, can include a series of treatments.

[0181] With regard to polypeptide formulations, featured herein is a method for treating osteoarthritis in a subject, which comprises contacting one or more cells in the subject with a first polypeptide, where the subject comprises a second polypeptide having one or more polymorphic variations associated with cancer, and where the first polypeptide comprises fewer polymorphic variations associated with cancer than the second polypeptide. The first and second polypeptides are encoded by a nucleic acid which comprises a nucleotide sequence in SEQ ID NO: 1-3; a nucleotide sequence which encodes a polypeptide consisting of an amino acid sequence encoded by a nucleotide sequence referenced in SEQ ID NO: 1-3; a nucleotide sequence which encodes a polypeptide that is 90% or more identical to an amino acid sequence encoded by a nucleotide sequence of SEQ ID NO: 1-3 and a nucleotide sequence 90% or more identical to a nucleotide sequence in SEQ ID NO: 1-3. The subject often is a human.

[0182] For antibodies, a dosage of 0.1 mg/kg of body weight (generally 10 mg/kg to 20 mg/kg) is often utilized. If the antibody is to act in the brain, a dosage of 50 mg/kg to 100 mg/kg is often appropriate. Generally, partially human antibodies and fully human antibodies have a longer half-life within the human body than other antibodies. Accordingly, lower dosages and less frequent administration is often possible. Modifications such as lipidation can be used to stabilize antibodies and to enhance uptake and tissue penetration (*e.g.*, into the brain). A method for lipidation of antibodies is described by Cruikshank *et al.*, *J. Acquired Immune Deficiency Syndromes and Human Retrovirology* 14:193 (1997).

[0183] Antibody conjugates can be used for modifying a given biological response, the drug moiety is not to be construed as limited to classical chemical therapeutic agents. For example, the drug moiety may be a protein or polypeptide possessing a desired biological activity. Such proteins may include, for example, a toxin such as abrin, ricin A, pseudomonas exotoxin, or diphtheria toxin; a polypeptide such as tumor necrosis factor, alpha-interferon, beta-interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator; or, biological response modifiers such as, for example, lymphokines, interleukin-1 (“IL-1”), interleukin-2 (“IL-2”), interleukin-6 (“IL-6”), granulocyte macrophage colony stimulating factor (“GM-CSF”), granulocyte colony stimulating factor (“G-CSF”), or other growth factors. Alternatively, an antibody can be conjugated to a second antibody to form an antibody heteroconjugate as described by Segal in U.S. Patent No. 4,676,980.

[0184] For compounds, exemplary doses include milligram or microgram amounts of the compound per kilogram of subject or sample weight, for example, about 1 microgram per kilogram to about 500 milligrams per kilogram, about 100 micrograms per kilogram to about 5 milligrams per kilogram, or about 1 microgram per kilogram to about 50 micrograms per kilogram. It is understood that appropriate doses of a small molecule depend upon the potency of the small molecule with respect to the expression or activity to be modulated. When one or more of these small molecules is to be administered to an animal (*e.g.*, a human) in order to modulate expression or activity of a polypeptide or nucleic acid described herein, a physician, veterinarian, or researcher may, for example, prescribe a relatively low dose at first, subsequently increasing the dose until an appropriate response is obtained. In addition, it is understood that the specific dose level for any particular animal subject will depend upon a variety of factors including the activity of the specific compound employed, the age, body weight, general health, gender, and diet of the subject, the time of administration, the route of administration, the rate of excretion, any drug combination, and the degree of expression or activity to be modulated.

[0185] With regard to nucleic acid formulations, gene therapy vectors can be delivered to a subject by, for example, intravenous injection, local administration (*see, e.g.*, U.S. Patent 5,328,470) or by stereotactic injection (*see e.g.*, Chen *et al.*, (1994) *Proc. Natl. Acad. Sci. USA* 91:3054-3057). Pharmaceutical preparations of gene therapy vectors can include a gene therapy vector in an acceptable

diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells (e.g., retroviral vectors) the pharmaceutical preparation can include one or more cells which produce the gene delivery system. Examples of gene delivery vectors are described herein.

Therapeutic Methods

[0186] A therapeutic formulation described above can be administered to a subject in need of a therapeutic for inducing a desired biological response.. Therapeutic formulations can be administered by any of the paths described herein. With regard to both prophylactic and therapeutic methods of treatment, such treatments may be specifically tailored or modified, based on knowledge obtained from pharmacogenomic analyses described herein.

[0187] As used herein, the term "treatment" is defined as the application or administration of a therapeutic formulation to a subject, or application or administration of a therapeutic agent to an isolated tissue or cell line from a subject with the purpose to cure, heal, alleviate, relieve, alter, remedy, ameliorate, improve or affect osteoarthritis, symptoms of osteoarthritis or a predisposition towards osteoarthritis. A therapeutic formulation includes, but is not limited to, small molecules, peptides, antibodies, ribozymes and antisense oligonucleotides. Administration of a therapeutic formulation can occur prior to the manifestation of symptoms characteristic of osteoarthritis, such that osteoarthritis is prevented or delayed in its progression. The appropriate therapeutic composition can be determined based on screening assays described herein.

[0188] As discussed, successful treatment of osteoarthritis can be brought about by techniques that serve to agonize target molecule expression or function, or alternatively, antagonize target molecule expression or function. These techniques include administration of modulators that include, but are not limited to, small organic or inorganic molecules; antibodies (including, for example, polyclonal, monoclonal, humanized, anti-idiotypic, chimeric or single chain antibodies, and Fab, F(ab')₂ and Fab expression library fragments, scFV molecules, and epitope-binding fragments thereof); and peptides, phosphopeptides, or polypeptides.

[0189] Further, antisense and ribozyme molecules that inhibit expression of the target gene can also be used to reduce the level of target gene expression, thus effectively reducing the level of target gene activity. Still further, triple helix molecules can be utilized in reducing the level of target gene activity. Antisense, ribozyme and triple helix molecules are discussed above. It is possible that the use of antisense, ribozyme, and/or triple helix molecules to reduce or inhibit mutant gene expression can also reduce or inhibit the transcription (triple helix) and/or translation (antisense, ribozyme) of mRNA produced by normal target gene alleles, such that the concentration of normal target gene product present can be lower than is necessary for a normal phenotype. In such cases, nucleic acid molecules that encode

and express target gene polypeptides exhibiting normal target gene activity can be introduced into cells via gene therapy method. Alternatively, in instances in that the target gene encodes an extracellular polypeptide, it can be preferable to co-administer normal target gene polypeptide into the cell or tissue in order to maintain the requisite level of cellular or tissue target gene activity.

[0190] Another method by which nucleic acid molecules may be utilized in treating or preventing osteoarthritis is use of aptamer molecules specific for target molecules. Aptamers are nucleic acid molecules having a tertiary structure which permits them to specifically bind to ligands (*see, e.g., Osborne, et al., Curr. Opin. Chem. Biol.* 1(1): 5-9 (1997); and Patel, D. J., *Curr. Opin. Chem. Biol.* Jun; 1(1): 32-46 (1997)).

[0191] Yet another method of utilizing nucleic acid molecules for osteoarthritis treatment is gene therapy, which can also be referred to as allele therapy. Provided herein is a gene therapy method for treating osteoarthritis in a subject, which comprises contacting one or more cells in the subject or from the subject with a nucleic acid having a first nucleotide sequence (e.g., the first nucleotide sequence is identical to or substantially identical to a nucleotide sequence of SEQ ID NO: 1-3). Genomic DNA in the subject comprises a second nucleotide sequence having one or more polymorphic variations associated with osteoarthritis (e.g., the second nucleotide sequence is identical to or substantially identical to a nucleotide sequence of SEQ ID NO: 1). The first and second nucleotide sequences typically are substantially identical to one another, and the first nucleotide sequence comprises fewer polymorphic variations associated with osteoarthritis than the second nucleotide sequence. The first nucleotide sequence may comprise a gene sequence that encodes a full-length polypeptide or a fragment thereof. The subject is often a human. Allele therapy methods often are utilized in conjunction with a method of first determining whether a subject has genomic DNA that includes polymorphic variants associated with osteoarthritis.

[0192] In another allele therapy embodiment, provided herein is a method which comprises contacting one or more cells in the subject or from the subject with a polypeptide encoded by a nucleic acid having a first nucleotide sequence (e.g., the first nucleotide sequence is identical to or substantially identical to the nucleotide sequence of SEQ ID NO: 1-3). Genomic DNA in the subject comprises a second nucleotide sequence having one or more polymorphic variations associated with osteoarthritis (e.g., the second nucleotide sequence is identical to or substantially identical to a nucleotide sequence of SEQ ID NO: 1). The first and second nucleotide sequences typically are substantially identical to one another, and the first nucleotide sequence comprises fewer polymorphic variations associated with osteoarthritis than the second nucleotide sequence. The first nucleotide sequence may comprise a gene sequence that encodes a full-length polypeptide or a fragment thereof. The subject is often a human. The method often comprises supplementing arthritis-associated *ADAMTS2* polypeptide with a non-arthritis-associated *ADAMTS2* polypeptide or fragment thereof, where the non-arthritis-associated form of

ADAMTS2 contains an isoleucine at position 245 of SEQ ID NO: 4 having enzymatic activity. The arthritis-associated *ADAMTS2* polypeptide sometimes contains a valine at position 245 of SEQ ID NO: 4 having an altered enzymatic activity varying from the non-arthritis-associated polypeptide.

[0193] In an embodiment, provided is a method of increasing the synthesis of procollagen II comprising providing or administering to individuals in need of increasing levels of type II collagen the pharmaceutical or physiologically acceptable composition comprising active human *ADAMTS2* protein or fragment thereof, where *ADAMTS2* polypeptide fragments having activity are selected from amino acids 252-1211, 253-1211, 254-1211, 255-1211, 256-1211, 257-1211, 258-1211, 259-1211 or 260-1211 of SEQ ID NO: 4, where it is understood that the active form of *ADAMTS2* does not contain the propeptide domain.

[0194] In another embodiment, provided herein is a method of increasing the synthesis of procollagen II comprising providing or administering to individuals in need of increasing levels of type II collagen the pharmaceutical or physiologically acceptable composition comprising an enzyme or molecule capable of cleaving *ADAMTS2* propeptide, e.g., a furin-type endopeptidase or N-ethylmaleimide described herein

[0195] For antibody-based therapies, antibodies can be generated that are both specific for target molecules and that reduce target molecule activity. Such antibodies may be administered in instances where antagonizing a target molecule function is appropriate for the treatment of osteoarthritis.

[0196] In circumstances where stimulating antibody production in an animal or a human subject by injection with a target molecule is harmful to the subject, it is possible to generate an immune response against the target molecule by use of anti-idiotypic antibodies (*see, e.g.*, Herlyn, *Ann. Med.*; 31(1): 66-78 (1999); and Bhattacharya-Chatterjee & Foon, *Cancer Treat. Res.*; 94: 51-68 (1998)). Introducing an anti-idiotypic antibody to a mammal or human subject often stimulates production of anti-anti-idiotypic antibodies, which typically are specific to the target molecule. Vaccines directed to osteoarthritis also may be generated in this fashion.

[0197] In instances where the target molecule is intracellular and whole antibodies are used, internalizing antibodies may be preferred. Lipofectin or liposomes can be used to deliver the antibody or a fragment of the Fab region that binds to the target antigen into cells. Where fragments of the antibody are used, the smallest inhibitory fragment that binds to the target antigen is preferred. For example, peptides having an amino acid sequence corresponding to the Fv region of the antibody can be used. Alternatively, single chain neutralizing antibodies that bind to intracellular target antigens can also be administered. Such single chain antibodies can be administered, for example, by expressing nucleotide sequences encoding single-chain antibodies within the target cell population (*see, e.g.*, Marasco *et al.*, *Proc. Natl. Acad. Sci. USA* 90: 7889-7893 (1993)).

[0198] Modulators can be administered to a patient at therapeutically effective doses to treat osteoarthritis. A therapeutically effective dose refers to an amount of the modulator sufficient to result in amelioration of symptoms of osteoarthritis. Toxicity and therapeutic efficacy of modulators can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the LD₅₀ (the dose lethal to 50% of the population) and the ED₅₀ (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD₅₀/ED₅₀. Modulators that exhibit large therapeutic indices are preferred. While modulators that exhibit toxic side effects can be used, care should be taken to design a delivery system that targets such molecules to the site of affected tissue in order to minimize potential damage to uninfected cells, thereby reducing side effects.

[0199] Data obtained from cell culture assays and animal studies can be used in formulating a range of dosages for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED₅₀ with little or no toxicity. The dosage can vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the methods described herein, the therapeutically effective dose can be estimated initially from cell culture assays. A dose can be formulated in animal models to achieve a circulating plasma concentration range that includes the IC₅₀ (*i.e.*, the concentration of the test compound that achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma can be measured, for example, by high performance liquid chromatography.

[0200] Another example of effective dose determination for an individual is the ability to directly assay levels of “free” and “bound” compound in the serum of the test subject. Such assays may utilize antibody mimics and/or “biosensors” that have been created through molecular imprinting techniques. Molecules that modulate target molecule activity are used as a template, or “imprinting molecule”, to spatially organize polymerizable monomers prior to their polymerization with catalytic reagents. The subsequent removal of the imprinted molecule leaves a polymer matrix which contains a repeated “negative image” of the compound and is able to selectively rebind the molecule under biological assay conditions. A detailed review of this technique can be seen in Ansell *et al.*, *Current Opinion in Biotechnology* 7: 89-94 (1996) and in Shea, *Trends in Polymer Science* 2: 166-173 (1994). Such “imprinted” affinity matrixes are amenable to ligand-binding assays, whereby the immobilized monoclonal antibody component is replaced by an appropriately imprinted matrix. An example of the use of such matrixes in this way can be seen in Vlatakis, *et al.*, *Nature* 361: 645-647 (1993). Through the use of isotope-labeling, the “free” concentration of compound which modulates target molecule expression or activity readily can be monitored and used in calculations of IC₅₀. Such “imprinted” affinity matrixes can also be designed to include fluorescent groups whose photon-emitting properties

measurably change upon local and selective binding of target compound. These changes readily can be assayed in real time using appropriate fiberoptic devices, in turn allowing the dose in a test subject to be quickly optimized based on its individual IC₅₀. An example of such a “biosensor” is discussed in Kriz *et al.*, *Analytical Chemistry* 67: 2142-2144 (1995).

[0201] The examples set forth below are intended to illustrate but not limit the invention.

Examples

[0202] In the following studies a group of subjects was selected according to specific parameters relating to osteoarthritis. Nucleic acid samples obtained from individuals in the study group were subjected to genetic analysis, which identified associations between osteoarthritis and a polymorphism in the *ADAMTS2* gene on chromosome five. The polymorphism was genotyped again in two replication cohorts consisting of individuals selected for OA. In addition, SNPs proximal to the incident polymorphism were identified and allelotyped in OA case and control pools. Methods are described for producing *ADAMTS2* polypeptide and *ADAMTS2* polypeptide variants *in vitro* or *in vivo*, *ADAMTS2* nucleic acids or polypeptides and variants thereof are utilized for screening test molecules for those that interact with *ADAMTS2* molecules. Test molecules identified as interactors with *ADAMTS2* molecules and *ADAMTS2* variants are further screened *in vivo* to determine whether they treat osteoarthritis.

Example 1

Samples and Pooling Strategies

Sample Selection

[0203] Blood samples were collected from individuals diagnosed with knee osteoarthritis, which were referred to as case samples. Also, blood samples were collected from individuals not diagnosed with knee osteoarthritis as gender and age-matched controls. A database was created that listed all phenotypic trait information gathered from individuals for each case and control sample. Genomic DNA was extracted from each of the blood samples for genetic analyses.

DNA Extraction from Blood Samples

[0204] Six to ten milliliters of whole blood was transferred to a 50 ml tube containing 27 ml of red cell lysis solution (RCL). The tube was inverted until the contents were mixed. Each tube was incubated for 10 minutes at room temperature and inverted once during the incubation. The tubes were then centrifuged for 20 minutes at 3000 x g and the supernatant was carefully poured off. 100-200 µl of residual liquid was left in the tube and was pipetted repeatedly to resuspend the pellet in the residual supernatant. White cell lysis solution (WCL) was added to the tube and pipetted repeatedly until

completely mixed. While no incubation was normally required, the solution was incubated at 37°C or room temperature if cell clumps were visible after mixing until the solution was homogeneous. 2 ml of protein precipitation was added to the cell lysate. The mixtures were vortexed vigorously at high speed for 20 sec to mix the protein precipitation solution uniformly with the cell lysate, and then centrifuged for 10 minutes at 3000 x g. The supernatant containing the DNA was then poured into a clean 15 ml tube, which contained 7 ml of 100% isopropanol. The samples were mixed by inverting the tubes gently until white threads of DNA were visible. Samples were centrifuged for 3 minutes at 2000 x g and the DNA was visible as a small white pellet. The supernatant was decanted and 5 ml of 70% ethanol was added to each tube. Each tube was inverted several times to wash the DNA pellet, and then centrifuged for 1 minute at 2000 x g. The ethanol was decanted and each tube was drained on clean absorbent paper. The DNA was dried in the tube by inversion for 10 minutes, and then 1000 µl of 1X TE was added. The size of each sample was estimated, and less TE buffer was added during the following DNA hydration step if the sample was smaller. The DNA was allowed to rehydrate overnight at room temperature, and DNA samples were stored at 2-8°C.

[0205] DNA was quantified by placing samples on a hematology mixer for at least 1 hour. DNA was serially diluted (typically 1:80, 1:160, 1:320, and 1:640 dilutions) so that it would be within the measurable range of standards. 125 µl of diluted DNA was transferred to a clear U-bottom microtitre plate, and 125 µl of 1X TE buffer was transferred into each well using a multichannel pipette. The DNA and 1X TE were mixed by repeated pipetting at least 15 times, and then the plates were sealed. 50 µl of diluted DNA was added to wells A5-H12 of a black flat bottom microtitre plate. Standards were inverted six times to mix them, and then 50 µl of 1X TE buffer was pipetted into well A1, 1000 ng/ml of standard was pipetted into well A2, 500 ng/ml of standard was pipetted into well A3, and 250 ng/ml of standard was pipetted into well A4. PicoGreen (Molecular Probes, Eugene, Oregon) was thawed and freshly diluted 1:200 according to the number of plates that were being measured. PicoGreen was vortexed and then 50µl was pipetted into all wells of the black plate with the diluted DNA. DNA and PicoGreen were mixed by pipetting repeatedly at least 10 times with the multichannel pipette. The plate was placed into a Fluoroskan Ascent Machine (microplate fluorometer produced by Labsystems) and the samples were allowed to incubate for 3 minutes before the machine was run using filter pairs 485 nm excitation and 538 nm emission wavelengths. Samples having measured DNA concentrations of greater than 450 ng/µl were re-measured for conformation. Samples having measured DNA concentrations of 20 ng/µl or less were re-measured for confirmation.

Pooling Strategies – Discovery Cohort

[0206] Samples were derived from the Nottingham knee OA family study (UK) where index cases were identified through a knee replacement registry. Siblings were approached and assessed with knee x-rays and assigned status as affected or unaffected. In all 1,157 individuals were available. In order to create same-sex pools of appropriate sizes, 335 unrelated female individuals with OA from the Nottingham OA sample were selected for the case pool. The control pool was made up of unrelated female individuals from the St. Thomas twin study (England) with normal knee x-rays and without other indications of OA, regardless of anatomical location, as well as lacking family history of OA. The St. Thomas twin study consists of Caucasian, female participants from the St. Thomas' Hospital, London, adult-twin registry, which is a voluntary registry of >4,000 twin pairs ranging from 18 to 76 years of age. The female case samples and female control samples are described further in Table 1 below.

[0207] A select set of samples from each group were utilized to generate pools, and one pool was created for each group. Each individual sample in a pool was represented by an equal amount of genomic DNA. For example, where 25 ng of genomic DNA was utilized in each PCR reaction and there were 200 individuals in each pool, each individual would provide 125 pg of genomic DNA. Inclusion or exclusion of samples for a pool was based upon the following criteria: the sample was derived from an individual characterized as Caucasian; the sample was derived from an individual of British paternal and maternal descent; case samples were derived from individuals diagnosed with specific knee osteoarthritis (OA) and were recruited from an OA knee replacement clinic. Control samples were derived from individuals free of OA, family history of OA, and rheumatoid arthritis. Also, sufficient genomic DNA was extracted from each blood sample for all allelotyping and genotyping reactions performed during the study. Phenotype information from each individual was collected and included age of the individual, gender, family history of OA, general medical information (e.g., height, weight, thyroid disease, diabetes, psoriasis, hysterectomy), joint history (previous and current symptoms, joint-related operations, age at onset of symptoms, date of primary diagnosis, age of individual as of primary diagnosis and order of involvement), and knee-related findings (crepitus, restricted passive movement, bony swelling/deformity). Additional knee information included knee history, current symptoms, any major knee injury, meniscectomy, knee replacement surgery, age of surgery, and treatment history (including hormone replace therapy (HRT)). Samples that met these criteria were added to appropriate pools based on disease status.

[0208] The selection process yielded the pools set forth in Table 1, which were used in the studies that follow:

TABLE 1

	Female case	Female control
Pool size (Number)	335	335
Pool Criteria (ex: case/control)	control	case
Mean Age (ex: years)	57.21	69.95

Example 2

Association of Polymorphic Variants with Osteoarthritis

[0209] A whole-genome screen was performed to identify particular SNPs associated with occurrence of osteoarthritis. As described in Example 1, two sets of samples were utilized, which included samples from female individuals having knee osteoarthritis (osteoarthritis cases), and samples from female individuals not having knee osteoarthritis (female controls). The initial screen of each pool was performed in an allelotyping study, in which certain samples in each group were pooled. By pooling DNA from each group, an allele frequency for each SNP in each group was calculated. These allele frequencies were then compared to one another. Particular SNPs were considered as being associated with osteoarthritis when allele frequency differences calculated between case and control pools were statistically significant. SNP disease association results obtained from the allelotyping study were then validated by genotyping each associated SNP across all samples from each pool. The results of the genotyping then were analyzed, allele frequencies for each group were calculated from the individual genotyping results, and a p-value was calculated to determine whether the case and control groups had statistically significant differences in allele frequencies for a particular SNP. When the genotyping results agreed with the original allelotyping results, the SNP disease association was considered validated at the genetic level.

SNP Panel Used for Genetic Analyses

[0210] A whole-genome SNP screen began with an initial screen of approximately 25,000 SNPs over each set of disease and control samples using a pooling approach. The pools studied in the screen are described in Example 1. The SNPs analyzed in this study were part of a set of 25,488 SNPs confirmed as being statistically polymorphic as each is characterized as having a minor allele frequency of greater than 10%. The SNPs in the set reside in genes or in close proximity to genes, and many reside

in gene exons. Specifically, SNPs in the set are located in exons, introns, and within 5,000 base-pairs upstream of a transcription start site of a gene. In addition, SNPs were selected according to the following criteria: they are located in ESTs; they are located in Locuslink or Ensembl genes; and they are located in Genomatix promoter predictions. SNPs in the set were also selected on the basis of even spacing across the genome, as depicted in Table 2.

[0211] A case-control study design using a whole genome association strategy involving approximately 28,000 single nucleotide polymorphisms (SNPs) was employed. Approximately 25,000 SNPs were evenly spaced in gene-based regions of the human genome with a median inter-marker distance of about 40,000 base pairs. Additionally, approximately 3,000 SNPs causing amino acid substitutions in genes described in the literature as candidates for various diseases were used. The case-control study samples were of female Caucasian origin (British paternal and maternal descent) 670 individuals were equally distributed in two groups: female controls and female cases. The whole genome association approach was first conducted on 2 DNA pools representing the 2 groups. Significant markers were confirmed by individual genotyping.

TABLE 2

<u>General Statistics</u>		<u>Spacing Statistics</u>	
Total # of SNPs	25,488	Median	37,058 bp
# of Exonic SNPs	>4,335 (17%)	Minimum*	1,000 bp
# SNPs with refSNP ID	20,776 (81%)	Maximum*	3,000,000 bp
Gene Coverage	>10,000	Mean	122,412 bp
Chromosome Coverage	All	Std Deviation	373,325 bp
<i>*Excludes outliers</i>			

Allelotyping and Genotyping Results

[0212] The genetic studies summarized above and described in more detail below identified an allelic variant in the ADAMTS2 gene that is associated with osteoarthritis.

Assay for Verifying, Allelotyping, and Genotyping SNPs

[0213] A MassARRAY™ system (Sequenom, Inc.) was utilized to perform SNP genotyping in a high-throughput fashion. This genotyping platform was complemented by a homogeneous, single-tube assay method (hME™ or homogeneous MassEXTEND™ (Sequenom, Inc.)) in which two genotyping primers anneal to and amplify a genomic target surrounding a polymorphic site of interest. A third primer (the MassEXTEND™ primer), which is complementary to the amplified target up to but not

including the polymorphism, was then enzymatically extended one or a few bases through the polymorphic site and then terminated.

[0214] For each polymorphism, SpectroDESIGNER™ software (Sequenom, Inc.) was used to generate a set of PCR primers and a MassEXTEND™ primer which were used to genotype the polymorphism. Other primer design software could be used or one of ordinary skill in the art could manually design primers based on his or her knowledge of the relevant factors and considerations in designing such primers. Table 3 shows PCR primers and Table 4 shows extension primers used for analyzing polymorphisms. The initial PCR amplification reaction was performed in a 5 µl total volume containing 1X PCR buffer with 1.5 mM MgCl₂ (Qiagen), 200 µM each of dATP, dGTP, dCTP, dTTP (Gibco-BRL), 2.5 ng of genomic DNA, 0.1 units of HotStar DNA polymerase (Qiagen), and 200 nM each of forward and reverse PCR primers specific for the polymorphic region of interest.

TABLE 3: PCR Primers

SNP Reference	Forward PCR primer	Reverse PCR primer
rs398829	ACGTTGGATGTAGTCATCGTCCGCAGCATG	ACGTTGGATGAAGACGGTGTCCTCTCCTTG

[0215] Samples were incubated at 95°C for 15 minutes, followed by 45 cycles of 95°C for 20 seconds, 56°C for 30 seconds, and 72°C for 1 minute, finishing with a 3 minute final extension at 72°C. Following amplification, shrimp alkaline phosphatase (SAP) (0.3 units in a 2 µl volume) (Amersham Pharmacia) was added to each reaction (total reaction volume was 7 µl) to remove any residual dNTPs that were not consumed in the PCR step. Samples were incubated for 20 minutes at 37°C, followed by 5 minutes at 85°C to denature the SAP.

[0216] Once the SAP reaction was complete, a primer extension reaction was initiated by adding a polymorphism-specific MassEXTEND™ primer cocktail to each sample. Each MassEXTEND™ cocktail included a specific combination of dideoxynucleotides (ddNTPs) and deoxynucleotides (dNTPs) used to distinguish polymorphic alleles from one another. Methods for verifying, allelotyping and genotyping SNPs are disclosed, for example, in U.S. Pat. No. 6,258,538, the content of which is hereby incorporated by reference. In Table 4, ddNTPs are shown and the fourth nucleotide not shown is the dNTP.

TABLE 4: Extension Primers

SNP Reference	Extend Probe	Termination Mix
rs398829	TGGCGTGCTCCTCTAGGA	ACG

[0217] The MassEXTEND™ reaction was performed in a total volume of 9 µl, with the addition of 1X ThermoSequenase buffer, 0.576 units of ThermoSequenase (Amersham Pharmacia), 600 nM MassEXTEND™ primer, 2 mM of ddATP and/or ddCTP and/or ddGTP and/or ddTTP, and 2 mM of dATP or dCTP or dGTP or dTTP. The deoxy nucleotide (dNTP) used in the assay normally was complementary to the nucleotide at the polymorphic site in the amplicon. Samples were incubated at 94°C for 2 minutes, followed by 55 cycles of 5 seconds at 94°C, 5 seconds at 52°C, and 5 seconds at 72°C.

[0218] Following incubation, samples were desalted by adding 16 µl of water (total reaction volume was 25 µl), 3 mg of SpectroCLEAN™ sample cleaning beads (Sequenom, Inc.) and allowed to incubate for 3 minutes with rotation. Samples were then robotically dispensed using a piezoelectric dispensing device (SpectroJET™ (Sequenom, Inc.)) onto either 96-spot or 384-spot silicon chips containing a matrix that crystallized each sample (SpectroCHIP™ (Sequenom, Inc.)). Subsequently, MALDI-TOF mass spectrometry (Biflex and Autoflex MALDI-TOF mass spectrometers (Bruker Daltonics) can be used) and SpectroTYPER RT™ software (Sequenom, Inc.) were used to analyze and interpret the SNP genotype for each sample.

Genetic Analysis

[0219] Minor allelic frequencies for the polymorphisms set forth in Table A were verified as being 10% or greater using the extension assay described above in a group of samples isolated from 92 individuals originating from the state of Utah in the United States, Venezuela and France (Coriell cell repositories).

[0220] Genotyping results are shown for female pools in Table 5. In Table 5, “AF” refers to allelic frequency; and “F case” and “F control” refer to female case and female control groups, respectively.

TABLE 5: Genotyping Results

SNP Reference	AF F case	AF F control	p-value
rs398829	G = 0.740 A = 0.260	G = 0.652 A = 0.348	0.0002

[0221] All of the single marker alleles set forth in Table A were considered validated, since the genotyping data agreed with the allelotyping data and each SNP significantly associated with osteoarthritis. Particularly significant associations with osteoarthritis are indicated by a calculated p-value of less than 0.05 for genotype results.

Example 3

Association of Polymorphic Variants with Osteoarthritis in Replication Cohorts

[0222] The single marker polymorphism set forth in Table A was genotyped again in two replication cohorts consisting of individuals selected for OA.

Sample Selection and Pooling Strategies – Replication Sample 1

[0223] A second case control sample (replication sample #1) was created by using 100 Caucasian female cases from Chingford, UK, and 148 unrelated female cases from the St. Thomas twin study. Cases were defined as having Kellgren-Lawrence (KL) scores of at least 2 in at least one knee x-ray. In addition, 199 male knee replacement cases from Nottingham were included. (For a cohort description, see the Nottingham description provided in Example 1). The control pool was made up of unrelated female individuals from the St. Thomas twin study (England) with normal knee x-rays and without other indications of OA, regardless of anatomical location, as well as lacking family history of OA. The St. Thomas twin study consists of Caucasian, female participants from the St. Thomas' Hospital, London, adult-twin registry, which is a voluntary registry of >4,000 twin pairs ranging from 18 to 76 years of age. The replication sample 1 cohort was used to replicate the initial results. Table 6 below summarizes the selected phenotype data collected from the case and control individuals.

TABLE 6

Phenotype	Female cases (n=248): median (range)/ (n,%)	Male cases (n=199): median (range)/ (n,%)	Female controls (n=313): mean (range)/ (n,%)
Age	59 (39- 73)	66 (45- 73)	55 (50- 72)
Height (cm)	162 (141- 178)	175 (152- 198)	162 (141- 176)
Weight (kg)	68 (51- 123)	86 (62- 127)	64 (40- 111)
Body mass index (kg/m ²)	26 (18- 44)	29 (21- 41)	24 (18- 46)
Kellgren-Lawrence* left knee	0 (63, 26%), 1 (20, 8%), 2 (105, 43%), 3 (58, 23%), 4 (1, 0%)	NA	NA
Kellgren-Lawrence* right knee	0 (43, 7%), 1 (18, 7%), 2 (127, 52%), 3 (57, 23%), 4 (1, 0%)	NA	NA
KL* >2 both knees	No (145, 59%), Yes (101, 41%)	NA	NA
KL* >2 either knee	No (0, 0%), Yes (248, 100%)	NA	NA

* 0: normal, 1: doubtful, 2: definite osteophyte (bony protuberance), 3: joint space narrowing (with or without osteophyte), 4: joint deformity

Sample Selection and Pooling Strategies – Replication Sample 2

[0224] A third case control sample (replication sample #2) was created by using individuals with symptoms of OA from Newfoundland, Canada. These individuals were recruited and examined by rheumatologists. Affected joints were x-rayed and a final diagnosis of definite or probable OA was made according to American College of Rheumatology criteria by a single rheumatologist to avoid any inter-examiner diagnosis variability. Controls were recruited from volunteers without any symptoms from the musculoskeletal system based on a normal joint exam performed by a rheumatologist. Only cases with a diagnosis of definite OA were included in the study. Only individuals of Caucasian origin were included. The cases consisted of 228 individuals with definite knee OA, 106 individuals with definite hip OA, and 74 individuals with hip OA.

TABLE 7

Phenotype	Case	Control
Age at Visit	62.7	52.5
Sex (Female/Male)	227/119	174/101
Knee OA Xray: No	35% (120)	80% (16)
	1% (4)	0% (0)
	64% (221)	20% (4)
Hip OA Xray: No	63% (215)	80% (16)
	2% (7)	0% (0)
	35% (121)	20% (4)

Assay for Verifying, Allelotyping, and Genotyping SNPs

[0225] Genotyping of the replication cohorts described in Tables 6 and 7 was performed using the same methods used for the original genotyping, as described herein. A MassARRAY™ system (Sequenom, Inc.) was utilized to perform SNP genotyping in a high-throughput fashion. This genotyping platform was complemented by a homogeneous, single-tube assay method (hMET™ or homogeneous MassEXTEND™ (Sequenom, Inc.)) in which two genotyping primers anneal to and amplify a genomic target surrounding a polymorphic site of interest. A third primer (the MassEXTEND™ primer), which is complementary to the amplified target up to but not including the polymorphism, was then enzymatically extended one or a few bases through the polymorphic site and then terminated.

[0226] For each polymorphism, SpectroDESIGNER™ software (Sequenom, Inc.) was used to generate a set of PCR primers and a MassEXTEND™ primer which where used to genotype the polymorphism. Other primer design software could be used or one of ordinary skill in the art could

manually design primers based on his or her knowledge of the relevant factors and considerations in designing such primers. Table 3 shows PCR primers and Table 4 shows extension probes used for analyzing (e.g., genotyping) polymorphisms in the replication cohorts. The initial PCR amplification reaction was performed in a 5 µl total volume containing 1X PCR buffer with 1.5 mM MgCl₂ (Qiagen), 200 µM each of dATP, dGTP, dCTP, dTTP (Gibco-BRL), 2.5 ng of genomic DNA, 0.1 units of HotStar DNA polymerase (Qiagen), and 200 nM each of forward and reverse PCR primers specific for the polymorphic region of interest.

[0227] Samples were incubated at 95°C for 15 minutes, followed by 45 cycles of 95°C for 20 seconds, 56°C for 30 seconds, and 72°C for 1 minute, finishing with a 3 minute final extension at 72°C. Following amplification, shrimp alkaline phosphatase (SAP) (0.3 units in a 2 µl volume) (Amersham Pharmacia) was added to each reaction (total reaction volume was 7 µl) to remove any residual dNTPs that were not consumed in the PCR step. Samples were incubated for 20 minutes at 37°C, followed by 5 minutes at 85°C to denature the SAP.

[0228] Once the SAP reaction was complete, a primer extension reaction was initiated by adding a polymorphism-specific MassEXTEND™ primer cocktail to each sample. Each MassEXTEND™ cocktail included a specific combination of dideoxynucleotides (ddNTPs) and deoxynucleotides (dNTPs) used to distinguish polymorphic alleles from one another. Methods for verifying, allelotyping and genotyping SNPs are disclosed, for example, in U.S. Pat. No. 6,258,538, the content of which is hereby incorporated by reference. In Table 7, ddNTPs are shown and the fourth nucleotide not shown is the dNTP.

[0229] The MassEXTEND™ reaction was performed in a total volume of 9 µl, with the addition of 1X ThermoSequenase buffer, 0.576 units of ThermoSequenase (Amersham Pharmacia), 600 nM MassEXTEND™ primer, 2 mM of ddATP and/or ddCTP and/or ddGTP and/or ddTTP, and 2 mM of dATP or dCTP or dGTP or dTTP. The deoxy nucleotide (dNTP) used in the assay normally was complementary to the nucleotide at the polymorphic site in the amplicon. Samples were incubated at 94°C for 2 minutes, followed by 55 cycles of 5 seconds at 94°C, 5 seconds at 52°C, and 5 seconds at 72°C.

[0230] Following incubation, samples were desalted by adding 16 µl of water (total reaction volume was 25 µl), 3 mg of SpectroCLEAN™ sample cleaning beads (Sequenom, Inc.) and allowed to incubate for 3 minutes with rotation. Samples were then robotically dispensed using a piezoelectric dispensing device (SpectroJET™ (Sequenom, Inc.)) onto either 96-spot or 384-spot silicon chips containing a matrix that crystallized each sample (SpectroCHIP™ (Sequenom, Inc.)). Subsequently, MALDI-TOF mass spectrometry (Biflex and Autoflex MALDI-TOF mass spectrometers (Bruker Daltonics) can be used) and

SpectroTYPER RT™ software (Sequenom, Inc.) were used to analyze and interpret the SNP genotype for each sample.

Genetic Analysis

[0231] Genotyping results for replication cohorts #1 and #2 are provided in Tables 8 and 9, respectively.

TABLE 8

rsID	Replication #1 (Mixed Male/Female cases and Female controls)				Meta-analysis Disc. + Rep #1 P-value
	AF OA Con	AF OA Cas	Delta	P-value	
rs398829	0.30	0.28	0.02	0.307	0.0260

TABLE 9

rsID	Replication #2 (Newfoundland) (Male/Female cases and controls)				Meta-analysis Disc. + Rep #2 Not Done
	AF OA Con	AF OA Cas	Delta	P-value	
rs398829	0.27	0.28	-0.013	0.627	

[0232] To combine the evidence for association from multiple sample collections, a meta-analysis procedure was employed. The allele frequencies were compared between cases and controls within the discovery sample, as well as within the replication cohort #1 using the DerSimian-Laird approach (DerSimonian, R. and N. Laird. 1986. Meta-analysis in clinical trials. Control Clin Trials 7: 177-188.)

[0233] The absence of a statistically significant association in one or more of the replication cohorts should not be interpreted as minimizing the value of the original finding. There are many reasons why a biologically derived association identified in a sample from one population would not replicate in a sample from another population. The most important reason is differences in population history. Due to bottlenecks and founder effects, there may be common disease predisposing alleles present in one population that are relatively rare in another, leading to a lack of association in the candidate region. Also, because common diseases such as arthritis-related disorders are the result of susceptibilities in many genes and many environmental risk factors, differences in population-specific genetic and environmental backgrounds could mask the effects of a biologically relevant allele. For these and other reasons, statistically strong results in the original, discovery sample that did not replicate in one or more of the replication samples may be further evaluated in additional replication cohorts and experimental systems.

Example 4
ADAMTS2 Region Proximal SNPs

[0234] It has been discovered that SNP rs398829 in *ADAMTS2* is associated with occurrence of osteoarthritis in subjects. This gene encodes a disintegrin and metalloproteinase with thrombospondin motifs-2 (*ADAMTS2*), which is a member of the *ADAMTS* protein family. Members of the family share several distinct protein modules, including a propeptide region, a metalloproteinase domain, a disintegrin-like domain, and a thrombospondin type 1 (TS) motif. *ADAMTS2* is involved in collagens 1, 2 and 5 N-terminal processing, (type II collagen is the major form in cartilage). Mutations in this gene cause Ehlers-Danlos syndrome type VIIC, a recessively inherited connective-tissue disorder that causes loose joints and fragile skin. Mild loss of function may exacerbate physical joint damage leading to a predisposition to OA and incorrectly processed collagen can act dominantly to inhibit self assembly of fibrils. Alternative splicing of the gene generates 2 transcript variants. The short transcript encodes a protein, which has no significant procollagen N-peptidase activity.

[0235] Two hundred-nine additional allelic variants proximal to rs398829 were identified and subsequently allelotyped in osteoarthritis case and control sample sets as described in Examples 1 and 2. The polymorphic variants are set forth in Table 10. The chromosome positions provided in column four of Table 10 are based on Genome “Build 34” of NCBI’s GenBank.

TABLE 10

dbSNP rs#	Chromosome	Position in SEQ ID NO: 1	Chromosome Position	Allele Variants
rs2278221	5	210	178695460	c/t
rs1650358	5	3608	178698858	c/g
rs1643818	5	3609	178698859	c/g
rs3733916	5	4318	178699568	c/t
rs1624933	5	5593	178700843	a/g
rs1624857	5	5629	178700879	c/t
rs1624832	5	5639	178700889	a/g
rs1624829	5	5640	178700890	c/t
rs2161171	5	8943	178704193	a/c
rs1530499	5	17968	178713218	a/g
rs888764	5	19887	178715137	a/g
rs873987	5	21034	178716284	a/g
rs4078699	5	21085	178716335	c/t
rs870311	5	21596	178716846	a/g
rs1643817	5	23379	178718629	a/c
rs1643816	5	23432	178718682	a/c
rs1650355	5	24007	178719257	a/c
rs888763	5	26121	178721371	a/g
rs1862212	5	26273	178721523	a/t

dbSNP rs#	Chromosome	Position in SEQ ID NO: 1	Chromosome Position	Allele Variants
rs1110514	5	26755	178722005	a/t
rs3797600	5	27411	178722661	c/t
rs3797602	5	27710	178722960	g/t
rs3797603	5	27842	178723092	c/t
rs3776819	5	28379	178723629	c/t
rs252076	5	29603	178724853	c/t
rs252075	5	31232	178726482	c/g
rs252074	5	31504	178726754	a/g
rs252068	5	32583	178727833	c/g
rs252069	5	32794	178728044	a/g
rs194040	5	32840	178728090	c/t
rs252070	5	33044	178728294	c/t
rs3797606	5	33150	178728400	a/c
rs171667	5	33218	178728468	a/g
rs187539	5	33513	178728763	c/t
rs3836834	5	33959	178729209	- <chem>/tatcaaactaccatgaaa</chem>
rs252071	5	34486	178729736	a/g
rs252072	5	36289	178731539	c/t
rs252073	5	36570	178731820	c/t
rs379589	5	38247	178733497	a/t
rs2052472	5	38477	178733727	a/c
rs2052471	5	38518	178733768	c/t
rs2052470	5	38529	178733779	c/t
rs2052469	5	38667	178733917	a/g
rs3797608	5	39781	178735031	c/t
rs3797609	5	39856	178735106	c/t
rs3822601	5	39927	178735177	c/t
rs153131	5	40506	178735756	a/g
rs751546	5	41869	178737119	c/g
rs2279979	5	42452	178737702	c/t
rs252060	5	44788	178740038	c/t
rs3797610	5	46059	178741309	a/c
rs194039	5	46846	178742096	a/g
rs168773	5	47712	178742962	a/t
rs252061	5	48796	178744046	c/t
rs187537	5	49441	178744691	c/g
rs252062	5	49602	178744852	a/t
rs2431255	5	49723	178744973	a/c
rs3797612	5	50050	178745300	c/t
rs3797613	5	50171	178745421	c/t
rs614114	5	50477	178745727	c/t
rs252063	5	50818	178746068	c/t
rs252064	5	50833	178746083	c/t
rs252065	5	50881	178746131	a/g
rs450502	5	50882	178746132	a/g
rs439252	5	51386	178746636	c/t

dbSNP rs#	Chromosome	Position in SEQ ID NO: 1	Chromosome Position	Allele Variants
rs252066	5	51534	178746784	c/t
rs457957	5	52317	178747567	a/g
rs3797614	5	52368	178747618	c/t
rs423552	5	52970	178748220	a/g
rs398829	5	53023	178748273	a/g
rs416646	5	53356	178748606	a/g
rs187450	5	53882	178749132	g/t
rs337807	5	54553	178749803	c/t
rs337806	5	55475	178750725	a/c
rs1396438	5	55530	178750780	a/g
rs1396437	5	55691	178750941	c/t
rs2411811	5	55848	178751098	a/c
rs2898813	5	55879	178751129	c/g
rs189256	5	56316	178751566	a/g
rs173072	5	56911	178752161	a/c
rs337805	5	57320	178752570	a/g
rs191415	5	57391	178752641	c/t
rs180045	5	57437	178752687	c/t
rs189255	5	57478	178752728	c/g
rs652766	5	57500	178752750	c/t
rs466750	5	59111	178754361	g/t
rs442406	5	59333	178754583	a/g
rs662407	5	59715	178754965	a/g
rs592971	5	59804	178755054	a/g
rs457187	5	59851	178755101	a/g
rs459490	5	59929	178755179	c/t
rs459668	5	60052	178755302	c/t
rs462646	5	60240	178755490	c/t
rs458272	5	60359	178755609	g/t
rs463455	5	60381	178755631	a/g
rs675880	5	60456	178755706	c/t
rs810617	5	60724	178755974	c/g
rs464156	5	60875	178756125	c/t
rs458083	5	60968	178756218	a/g
rs467333	5	60978	178756228	c/g
rs465381	5	60998	178756248	c/t
rs466363	5	61557	178756807	c/t
rs2457099	5	62091	178757341	c/t
rs463901	5	62645	178757895	c/t
rs465621	5	62943	178758193	a/c
rs463724	5	63131	178758381	a/t
rs465242	5	63145	178758395	g/t
rs467419	5	63406	178758656	a/g
rs456135	5	63427	178758677	c/g
rs464536	5	63554	178758804	c/t
rs461898	5	63661	178758911	a/g
rs389558	5	64093	178759343	a/g

dbSNP rs#	Chromosome	Position in SEQ ID NO: 1	Chromosome Position	Allele Variants
rs466752	5	64153	178759403	c/t
rs455655	5	64409	178759659	c/g
rs463435	5	64544	178759794	c/t
rs2174971	5	65257	178760507	c/t
rs1979979	5	65626	178760876	a/g
rs411804	5	65739	178760989	a/g
rs1623885	5	66392	178761642	c/t
rs1643811	5	66720	178761970	c/t
rs434430	5	69177	178764427	a/t
rs187538	5	69336	178764586	g/t
rs252067	5	69636	178764886	a/g
rs459319	5	69823	178765073	a/g
rs467289	5	69928	178765178	c/t
rs462644	5	70547	178765797	c/t
rs458752	5	70633	178765883	c/t
rs708320	5	71805	178767055	a/c
rs457954	5	72181	178767431	c/g
rs2411810	5	72200	178767450	c/t
rs3084687	5	72474	178767724	-/at
rs69638	5	72567	178767817	c/g
rs455452	5	72973	178768223	a/g
rs464850	5	73468	178768718	a/g
rs431472	5	73889	178769139	a/g
rs2411809	5	75730	178770980	c/t
rs2457094	5	75970	178771220	a/g
rs2457095	5	76114	178771364	a/g
rs2261740	5	76342	178771592	c/t
rs1109180	5	76449	178771699	a/g
rs1109179	5	76465	178771715	c/t
rs1109178	5	76791	178772041	a/c
rs456909	5	78042	178773292	a/g
rs469124	5	80758	178776008	a/g
rs468039	5	80778	178776028	c/t
rs467017	5	81356	178776606	a/c
rs469290	5	81576	178776826	a/g
rs469090	5	81689	178776939	c/t
rs469568	5	81759	178777009	g/t
rs468386	5	81950	178777200	c/g
rs469349	5	82562	178777812	a/c
rs469099	5	83591	178778841	c/t
rs456868	5	83700	178778950	a/g
rs465389	5	83821	178779071	c/g
rs463892	5	83842	178779092	c/g
rs468548	5	83923	178779173	g/t
rs654612	5	83929	178779179	a/c
rs468542	5	84021	178779271	c/g
rs469262	5	84175	178779425	c/t

dbSNP rs#	Chromosome	Position in SEQ ID NO: 1	Chromosome Position	Allele Variants
rs708323	5	84417	178779667	a/g
rs469089	5	84747	178779997	c/g
rs469396	5	85746	178780996	c/g
rs468723	5	86129	178781379	c/t
rs467604	5	86335	178781585	a/g
rs338874	5	87315	178782565	c/g
rs338875	5	87648	178782898	a/g
rs1385803	5	87764	178783014	a/c
rs1385804	5	87770	178783020	c/g
rs338876	5	88221	178783471	c/t
rs189803	5	90474	178785724	a/c
rs452215	5	91148	178786398	g/t
rs641170	5	91150	178786400	g/t
rs584398	5	91160	178786410	g/t
rs385330	5	91733	178786983	c/t
rs429538	5	91772	178787022	a/c
rs371229	5	91785	178787035	c/t
rs460874	5	93140	178788390	a/t
rs646121	5	93148	178788398	a/t
rs468262	5	96080	178791330	a/g
rs467863	5	96157	178791407	c/g
rs191434	5	96313	178791563	a/c
rs2054782	5	96759	178792009	c/t
rs468499	5	97026	178792276	a/c
rs180287	5	97320	178792570	c/g
rs338877	5	97732	178792982	a/t
rs650665	5	98713	178793963	c/g
rs193419	5	99707	178794957	a/c
rs180288	5	99959	178795209	c/g
rs186834	5	100009	178795259	a/g
rs189266	5	100020	178795270	c/g
rs189267	5	100065	178795315	a/c
rs170937	5	100086	178795336	c/g
rs463263	5	101270	178796520	c/g
rs463262	5	101276	178796526	g/t
rs460454	5	101371	178796621	c/t
rs460455	5	101376	178796626	c/g
rs460505	5	101439	178796689	c/t
rs931316	5	101820	178797070	c/t
rs463431	5	102392	178797642	c/g
rs461542	5	102602	178797852	a/g
rs463557	5	102604	178797854	a/c
rs191453	5	102896	178798146	c/t
rs2271212	5	189104	178884354	c/t
rs462009	5	189134	178884384	c/t
rs2271211	5	189205	178884455	a/g
rs396474	5	Not mapped	Not mapped	a/c

dbSNP rs#	Chromosome	Position in SEQ ID NO: 1	Chromosome Position	Allele Variants
rs428901	5	Not mapped	Not mapped	a/t
rs452300	5	Not mapped	Not mapped	g/t
rs670256	5	Not mapped	Not mapped	g/t

Assay for Verifying and Allelotyping SNPs

[0236] The methods used to verify and allelotype the 209 proximal SNPs of Table 10 are the same methods described in Examples 1 and 2 herein. The primers and probes used in these assays are provided in Table 11 and Table 12, respectively.

TABLE 11

dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs2278221	ACGTTGGATGTCTCATGGGCCACCACAAAC	ACGTTGGATGTATGCTCCTGTCACCGGCAT
rs1650358	ACGTTGGATGTGGATGGCTCCATGTTCTG	ACGTTGGATGAAGTGTGGATTACAGGTG
rs1643818	ACGTTGGATGCTGGGATTACAGGTGTGAAC	ACGTTGGATGTGGATGGCTCCATGTTCTG
rs3733916	ACGTTGGATGCCGAGCAGGCTGTAGTGTG	ACGTTGGATGCTTGATACCACCTGGAACAG
rs1624933	ACGTTGGATGAGGCTGGCTCAAACCTCTG	ACGTTGGATGTAACAAAAAGTTGGCCGTGC
rs1624857	ACGTTGGATGTGAGGTCAAGGAGTTGAGAC	ACGTTGGATGGCCACCAAGCCAGACTAAGT
rs1624832	ACGTTGGATGTGAGGTCAAGGAGTTGAGAC	ACGTTGGATGGCCACCAAGCCAGACTAAGT
rs1624829	ACGTTGGATGTGAGGTCAAGGAGTTGAGAC	ACGTTGGATGGCCACCAAGCCAGACTAAGT
rs2161171	ACGTTGGATGCCCGTCACCACTTTATTCC	ACGTTGGATGAGAGTGGATCCAGTCTGCAG
rs1530499	ACGTTGGATGACTCCAAGATTCCCATTTC	ACGTTGGATGTTCTGTGTTCCACCTATGG
rs888764	ACGTTGGATGTAGTTGAATGTTGTATTGGC	ACGTTGGATGACCGTGATAAACACAGAATG
rs873987	ACGTTGGATGGCTGTTAACATGTGTCGGG	ACGTTGGATGATTGGCCACATCACCAAGAC
rs4078699	ACGTTGGATGGTACCGTGGATTCTTTAGG	ACGTTGGATGGTATTGGAAAAGAGCAGAGAC
rs870311	ACGTTGGATGTCAGGGCTCCAGTGTGAAG	ACGTTGGATGAAAAGGAGGAGTGCCCTGTG
rs1643817	ACGTTGGATGATGGGAAACTCCTGGCCTG	ACGTTGGATGAAAATGCAAGCCGCCACCTG
rs1643816	ACGTTGGATGTTCTCCCTTAGCCC	ACGTTGGATGTTGGCATGAGAGATGGACAG
rs1650355	ACGTTGGATGTCACAGCAACAAAACAAA	ACGTTGGATGTTAACATGGTCAGAGGGTTG
rs888763	ACGTTGGATGAAGAGGAAGAGACATACCAG	ACGTTGGATGAACAACATGGACTCAGGCTG
rs1862212	ACGTTGGATGGGCCACATTAAAACAAGGG	ACGTTGGATGTCCTGAGGTTCTATAAG
rs1110514	ACGTTGGATGTGCCAACGTTCCATGTTAG	ACGTTGGATGATCACTGTAGCCCCCTTCCTG
rs3797600	ACGTTGGATGCCCTCCTGTCACCTCCTTG	ACGTTGGATGGGAAGTGACTGCTGAGCTG
rs3797602	ACGTTGGATGAGAACAGGGACTGGCTGT	ACGTTGGATGAGCAGGCTCCGGAAAGTATG
rs3797603	ACGTTGGATGCACCCATCCATCATGATGTC	ACGTTGGATGTGCTACCTCAAAACAGTGGG
rs3776819	ACGTTGGATGCAAGCACCATTCCATTGCAC	ACGTTGGATGAAATGAGGATTGCAGTCCCC
rs252076	ACGTTGGATGACTTCTGACTTCAGGTGATC	ACGTTGGATGTATAGGAACGAAAAGAACCC
rs252075	ACGTTGGATGTGGGAGCATTGCAAGGCATG	ACGTTGGATGAAGCCTCAGATGGTTCCGGAG
rs252074	ACGTTGGATGTTGCGATGGCCTCTGGCT	ACGTTGGATGAAGTTGAGGGCTCCGGAGCA
rs252068	ACGTTGGATGGGGTAGGAAGGGTTAAGC	ACGTTGGATGGCAGCCCCCTCAATTCTTAG
rs252069	ACGTTGGATGTGCCATTCCCTGTTATTCC	ACGTTGGATGTTGGACTTGCCGTGCAACT
rs194040	ACGTTGGATGTGCCATTCCCTGTTATTCC	ACGTTGGATGTTGGACTTGCCGTGCAACT
rs252070	ACGTTGGATGCTCAAGGACATTGTCCTGG	ACGTTGGATGGGAGAAGCAGCTCCCTTC

dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs3797606	ACGTTGGATGGTTCCCCAAACAAGAGAGC	ACGTTGGATGGGAAATGTTCAAAGCCGCAG
rs171667	ACGTTGGATGGGAACACATTGTAATGCG	ACGTTGGATGCCCTCCTCATTGTCTATTCC
rs187539	ACGTTGGATGAGCCACCCCAACCTTCAGGA	ACGTTGGATGTTGCTCCTGGACATGGTTT
rs3836834	ACGTTGGATGAAGAACGTGACTCTGCTC	ACGTTGGATGTAGTAATTCTGATCCTGGCC
rs252071	ACGTTGGATGGCTCAACCTGAAACAACCC	ACGTTGGATGGGATATTCCCACCTGAG
rs252072	ACGTTGGATGTTGTTCCCCAAGGCGACG	ACGTTGGATGTGTGTTCCAGAGCTGGAG
rs252073	ACGTTGGATGGGAAGGCCGAGAAAAGTC	ACGTTGGATGACAAGCTCAGCAGAGTCCA
rs379589	ACGTTGGATGAAACACGGAGTACTGAGCA	ACGTTGGATGTTGTTAGCTGTCTGTCCGTC
rs2052472	ACGTTGGATGAACCAGCTCAAGGATCACCC	ACGTTGGATGAAAGGAGACGGTCAGCTGTC
rs2052471	ACGTTGGATGACAGCTGACCGTCTCCTTTG	ACGTTGGATGCCCGTCTGGACAAGCTTT
rs2052470	ACGTTGGATGACAGCTGACCGTCTCCTTTG	ACGTTGGATGCCCGTCTGGACAAGCTTT
rs2052469	ACGTTGGATGAGGGAAAGATATCGCACGCG	ACGTTGGATGAGTGAACAAC TGCTCGCCTC
rs3797608	ACGTTGGATGTGCTTGCTTGGCTCTGC	ACGTTGGATGTGCACTAAGGGAGTGAGTGG
rs3797609	ACGTTGGATGTGCGAGAACCCAAGGCAAAGC	ACGTTGGATGACAGCATTTGGAGTCCCCTG
rs3822601	ACGTTGGATGAGGTCA GTGAGGCCTGAGAT	ACGTTGGATGTGCTGGCCTGAAGATCGAG
rs153131	ACGTTGGATGTAATCACGTGCTGTGATCCC	ACGTTGGATGAGCTGCTCTAGTCATGTC
rs751546	ACGTTGGATGTCCTGCTCTGCCGTTCTACA	ACGTTGGATGATCAGCTAAAGGACCGGTG
rs2279979	ACGTTGGATGTATTGCTACCAAGAACACGTA	ACGTTGGATGAAAAGGGGCCACTTCAGGG
rs252060	ACGTTGGATGTGGCCAGAGGCCGTGTTTC	ACGTTGGATGCCCAATCCCCTCTCATATG
rs3797610	ACGTTGGATGAAAAGCTTCTCCCTGGGTG	ACGTTGGATGCAAGTAGGGCAGAAACTCAG
rs194039	ACGTTGGATGAAAGTGTGGGATTACAGGC	ACGTTGGATGTGCTGGGAGAACATTAC
rs168773	ACGTTGGATGTGGTGCACTGGAGATATCAC	ACGTTGGATGGATCCCTATCTACCTCTTC
rs252061	ACGTTGGATGTGTCACACTCCTCTTGTAAAG	ACGTTGGATGCTGTGCTTCCATGCTTTGC
rs187537	ACGTTGGATGCCAGGATGTGATGCTAAGTG	ACGTTGGATGGTACCTCGCATAAGTGGATC
rs252062	ACGTTGGATGAAGCACATTGATGTGGCTGG	ACGTTGGATGCTGAAACTCAATGGGCACAG
rs2431255	ACGTTGGATGGGTGAAGACGGTACTTATG	ACGTTGGATGCTGGTGTGCTTGAAGAAC
rs3797612	ACGTTGGATGAGTGAGGACGCAGGGATTTC	ACGTTGGATGAGCGTGGCGAGGGAGATAA
rs3797613	ACGTTGGATGATCAGAGGCAGAGACCCCCC	ACGTTGGATGGGTGTCCTGCAGAGGC
rs614114	ACGTTGGATGGGTGGAGGATGTCTAGAAC	ACGTTGGATGGCTGGATCACTAGGGTTG
rs252063	ACGTTGGATGTTGAAATTACAGTCCGATGG	ACGTTGGATGCTGAGAGACTGAAAGCACA
rs252064	ACGTTGGATGCTGAGAGACTGAAAGCACA	ACGTTGGATGTTGAAATTACAGTCCGATGG
rs252065	ACGTTGGATGAAA ACTAAGGCTCAGAGGAC	ACGTTGGATGTGGCTGGATTACAGTCC
rs450502	ACGTTGGATGATGAGAAAACCAAGGCTCAG	ACGTTGGATGCTGGCTGGATTACAGTC
rs439252	ACGTTGGATGATCTCTGACCTCGTGTATCC	ACGTTGGATGTCTAGAACGGCCGGGTGC
rs252066	ACGTTGGATGTTCTCTTGACCGGTCTT	ACGTTGGATGTAACGAATTCTGCCGATG
rs457957	ACGTTGGATGTTCA CGTG CATTAGAGCGAG	ACGTTGGATGAATTCCCTCCCCAATTCTC
rs3797614	ACGTTGGATGACTGCGAGCTTAAGGAGGG	ACGTTGGATGCCAAACAGAACGGCCCTTTC
rs423552	ACGTTGGATGCCAGGACCTCGATGTTGTAG	ACGTTGGATGATCCTAGAGGAGCACGCCAAC
rs398829	ACGTTGGATGTAGTCATCGTCCGCAGCATG	ACGTTGGATGAAGACGGTGTCTCTCTTG
rs416646	ACGTTGGATGGCTGGCTCTCACAGTCTC	ACGTTGGATGAGACAGGCACCTCTGTGACTT
rs187450	ACGTTGGATGAGAAGGCAGGGACGATATCC	ACGTTGGATGACCAAGATGAACCCCTCTGT
rs337807	ACGTTGGATGTCACCCAGTGCTGACAGCAG	ACGTTGGATGATGCTGGGATGCCATGGGT
rs337806	ACGTTGGATGAATTAAAGAGATGGGCCACC	ACGTTGGATGCCCTGTGTGTTGTCTCC
rs1396438	ACGTTGGATGTACCTCTGGTGCAGAATG	ACGTTGGATGCCCTGGAGACAAAACACACAG
rs1396437	ACGTTGGATGAAAAACTCGCCTGCTCGG	ACGTTGGATGTCCAGACATTCCCCGTAGGA
rs2411811	ACGTTGGATGGAGGGATGCTCTAGAACATA	ACGTTGGATGCTGAATTTCACCTGAAATGG
rs2898813	ACGTTGGATGTCTCACCCACTTGCCTT	ACGTTGGATGATCGTATAATTGGGGTG

dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs189256	ACGTTGGATGCTCCCTATAAGCAAGGCTCTA	ACGTTGGATGTTAACCCAGGCCATGAAGAG
rs173072	ACGTTGGATGAGCTGGAGATCTCTTGCTC	ACGTTGGATGCTAAAACAGGATGGCTCTGG
rs337805	ACGTTGGATGGAAACAAACCAAGGAGCAGG	ACGTTGGATGATGTGGACAACGTTGGACTC
rs191415	ACGTTGGATGAATTACATGACTCGGACAAG	ACGTTGGATGTGCTGGTAAGTACAGAAGG
rs180045	ACGTTGGATGGTCCCAGGTTCTGTTCTC	ACGTTGGATGTGTACTTCACCAGCACTGAG
rs189255	ACGTTGGATGAGGTTGCAGACTCAGTCCC	ACGTTGGATGGGTGATTGCGGGAATGAG
rs652766	ACGTTGGATGGGTGATTGCGGGAATGAG	ACGTTGGATGACCATCCCACGATGCTCCC
rs466750	ACGTTGGATGTATCTCTTAAATGCCTTG	ACGTTGGATGTGACCAGGAGGAGTAAAAC
rs442406	ACGTTGGATGTGACAAGGTACGTGTTCTG	ACGTTGGATGCCAGACAAGTCTGATACAGC
rs662407	ACGTTGGATGCCACAGTCACCATTACTGAG	ACGTTGGATGCTTGAGCCATGAGTGGAAATG
rs592971	ACGTTGGATGGGAAGCATTCTTACTGC	ACGTTGGATGATTCCATCTCATGGCTCAAG
rs457187	ACGTTGGATGTGAGATGAGGAGTATCTG	ACGTTGGATGGCAGTCAAAGAAATGCTTCC
rs459490	ACGTTGGATGACAGATACTCCTCATCTCAC	ACGTTGGATGGGAGTTTGTGTTATAGC
rs459668	ACGTTGGATGGCTTCATTAAGTGGTCTTC	ACGTTGGATGTGAATGTTCAACGACTACAC
rs462646	ACGTTGGATGCAATTATTGACGGAGATTA	ACGTTGGATGCTCCTCCAAATGAATCAAGAA
rs458272	ACGTTGGATGATGCCCTCATTGTCATT	ACGTTGGATGCCAACAAAGTATTCCAAC
rs463455	ACGTTGGATGATGCCCTCATTGTCATT	ACGTTGGATGCCAACAAAGTATTCCAAC
rs675880	ACGTTGGATGCAGCTCCATTGATCTGTTTC	ACGTTGGATGAAGAATGACAATGAGGAGGC
rs810617	ACGTTGGATGTGATCTCAGCTTACACAGC	ACGTTGGATGATGCCTGTAATCCCAGCTAC
rs464156	ACGTTGGATGCAGATCCAAGAATATGTGGG	ACGTTGGATGTTCTAGAAAGGAGGCCAATC
rs458083	ACGTTGGATGTGTTCTTCTCCCCCTCCTG	ACGTTGGATGTGGCTCCTTCTAGAATCCC
rs467333	ACGTTGGATGCTTGTATTCTCCCCCTCC	ACGTTGGATGTTGGCTCCTTCTAGAATCC
rs465381	ACGTTGGATGACTTGCCTCATGTTCCAG	ACGTTGGATGACAAGCCTCTAAGGATAGGG
rs466363	ACGTTGGATGAAGTGACCCCTGAGGTGATGG	ACGTTGGATGTGAAGACAGTCACCCGTG
rs2457099	ACGTTGGATGTCTCCTTACACTGCCAGCGT	ACGTTGGATGCACTGTATTGACTTGGAC
rs463901	ACGTTGGATGAGAGTGCCAAGTGCAAAAGG	ACGTTGGATGTGCTTGCCTGCTGTGATCC
rs465621	ACGTTGGATGGGAAGTCATGGAAGTGCTAG	ACGTTGGATGAAAGAGCCCTAGGCTTGGAA
rs463724	ACGTTGGATGAGTGTGCCCTGTCGCCCTCA	ACGTTGGATGAAGGGCAGATGGCACACTTG
rs465242	ACGTTGGATGAGTGTGCCCTGTCGCCCTCA	ACGTTGGATGAAGGGCAGATGGCACACTTG
rs467419	ACGTTGGATGAGTCCCCAAACGTAAGTCC	ACGTTGGATGAGTCTAATCCCTGAGCCTC
rs456135	ACGTTGGATGAGTCTAATCCCTGAGCCTC	ACGTTGGATGACGTAAGTCTTAATGACCGC
rs464536	ACGTTGGATGTGCTCCAGGCTTGGCCTC	ACGTTGGATGAATTAGACTAAGGCCATGATG
rs461898	ACGTTGGATGGGAATACACAGCCACAGAG	ACGTTGGATGAGGTCAACGGAACAGGTC
rs389558	ACGTTGGATGCCAGTCCTGACAGTTCTA	ACGTTGGATGTTTCTCCCTGAAGCATGG
rs466752	ACGTTGGATGGCCTCTCCCTTAGTGC	ACGTTGGATGAGTCCTGACAGTTCTAAA
rs455655	ACGTTGGATGCTATTGCAACCCATATGGC	ACGTTGGATGAACACACAGCATCAGGTTCC
rs463435	ACGTTGGATGTTCAGCCATAGCTGGATTG	ACGTTGGATGCTCTGCTGGAAAATGTGAC
rs2174971	ACGTTGGATGAACACAACCTCCCTCGTC	ACGTTGGATGTGAATCCTGGAGGTGAGTG
rs1979979	ACGTTGGATGTGGCTGTCAGCACCCACTT	ACGTTGGATGCCAAAGGAAGGGAGAATTG
rs411804	ACGTTGGATGCAGATGACAGGCGGAAAATC	ACGTTGGATGAGGCTCCAGATGATGTCCA
rs1623885	ACGTTGGATGAATCAGCTAGGAAGAGCCTG	ACGTTGGATGTTCTGACCCCTCTAGGTCAG
rs1643811	ACGTTGGATGCAGGGCCCTGGTACTTCAG	ACGTTGGATGCATGGTGGTATTGCACCTG
rs434430	ACGTTGGATGTCAGGAGTTCACTGTAGAG	ACGTTGGATGCACATGCATACATTCATCAC
rs187538	ACGTTGGATGACATGGGCTTGGCAAATG	ACGTTGGATGCACCTGCTCAGAAGTAGCAT
rs252067	ACGTTGGATGAGAATTGCTGTGGTGTGAGG	ACGTTGGATGTTTCTTGGAGCTGCGC
rs459319	ACGTTGGATGCCATCTCTGACCTAGACA	ACGTTGGATGGCTCCAAGGAAAATTGGGAG
rs467289	ACGTTGGATGGCCCTTGGCTTGTCTT	ACGTTGGATGAGGCAGTGTGCCCTCTCATC

dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs462644	ACGTTGGATGATGTGGGTGAGCCCTTG	ACGTTGGATGTAACACTCAGCACGCCAG
rs458752	ACGTTGGATGCACCCACATCATGTGCGCTT	ACGTTGGATGCCCTCTCTACCCAGCACTT
rs708320	ACGTTGGATGAAACCAGCCTGGCTAACATG	ACGTTGGATGACAGGTGCCTGCTATCATA
rs457954	ACGTTGGATGAACCAGACCTGACTGATGG	ACGTTGGATGCCTCATACAAGTAGCCAAGG
rs2411810	ACGTTGGATGGCTTAACCAGACCTGACTG	ACGTTGGATGAGTGTAAAGGATATCCACGGC
rs3084687	ACGTTGGATGATCCCTGAGCCAGAGATT	ACGTTGGATGATGTCCTGTGCACACACAAG
rs69638	ACGTTGGATGTGCTCATTGCTGCTCATC	ACGTTGGATGAGAAGAAAGGTGTGCAGTGG
rs455452	ACGTTGGATGAGTGTGAGCCTGCTGG	ACGTTGGATGTCAGGTTCCCTCTGTGTC
rs464850	ACGTTGGATGTCTCTGTGCTCCAGACCA	ACGTTGGATGTGGGCTGAGATTCTGTGGG
rs431472	ACGTTGGATGAACCAGTGTGGGTGTGAAGC	ACGTTGGATGAGAGACTGCATCAGGCAGGA
rs2411809	ACGTTGGATGAGCGCATAAGTGACCAACCAG	ACGTTGGATGGCACTCACAGGGCATTGATG
rs2457094	ACGTTGGATGTTACTGTCACCTGGGTCTC	ACGTTGGATGGGAAGTCTGTATAGACGCAG
rs2457095	ACGTTGGATGTTATCAAGGCCTGCGCAGTG	ACGTTGGATGACTCCTGACCTCAGGCAATC
rs2261740	ACGTTGGATGATCGTGCCACTGCACTCCAG	ACGTTGGATGTCATCTTTGGTAGCCCCCCC
rs1109180	ACGTTGGATGCCAGGCCTGTATTGCACATC	ACGTTGGATGAGAATGCGTGTGCATGTGGG
rs1109179	ACGTTGGATGTAAATGGTATGCAGACCCC	ACGTTGGATGGAGTGCCGTATTGTCCTTC
rs1109178	ACGTTGGATGGCAAACAACAACAGCAACAG	ACGTTGGATGAAGTGTGGATTGTGCAGAC
rs456909	ACGTTGGATGTAGCTGCTTCATCTGTAAAG	ACGTTGGATGGGACTTTACCGATCTACTC
rs469124	ACGTTGGATGACTGGACACACATAGGCTG	ACGTTGGATGTGAAAATGCTAGGGTGTG
rs468039	ACGTTGGATGTGAAATGCTCAGGGTGTG	ACGTTGGATGAGGACTTGGACACACATAGG
rs467017	ACGTTGGATGGCTAGCTGCCACTAAACAG	ACGTTGGATGATGTGCCAAGAGGCTTGAG
rs469290	ACGTTGGATGTGCCCTTGTGCTCAGAG	ACGTTGGATGTCCTCTGTGCTGTGTTGG
rs469090	ACGTTGGATGACTGTCTTCAGGTGCTTGG	ACGTTGGATGGATGGTAGTCTCCTGGTT
rs469568	ACGTTGGATGAGCACCTCTGGCTTCATTG	ACGTTGGATGATTCAACCAGGAAATCCCAAC
rs468386	ACGTTGGATGTAATCCCAGCCCTTGGAAAG	ACGTTGGATGTATGGAGACAGGGTTTAC
rs469349	ACGTTGGATGTTAGAGACAGAGTCTCACTC	ACGTTGGATGTTGATCCCAGGAGTTCAAGG
rs469099	ACGTTGGATGTTGGAGCTGCTCTAGTTCTC	ACGTTGGATGTGAAAACCAGGGACTCAGCTC
rs456868	ACGTTGGATGACAGAGCAGGGAGCTCGGGT	ACGTTGGATGATTCAACCCAGCTACTGTG
rs465389	ACGTTGGATGAGGCTTGTAGACAGCTCCC	ACGTTGGATGTGCCAGTGCTCTGAGTATGC
rs463892	ACGTTGGATGAGGCTTGTAGACAGCTCCC	ACGTTGGATGTGCCAGTGCTCTGAGTATGC
rs468548	ACGTTGGATGACTGGAAAGGGAACATGCAAG	ACGTTGGATGCCCTGGATGCCCTTATAGAC
rs654612	ACGTTGGATGACTGGAAAGGGAACATGCAAG	ACGTTGGATGTGGATGCCCTTCTAGACAC
rs468542	ACGTTGGATGGCCTCATTTCCTTCTCAC	ACGTTGGATGTGCTAGAAAGGGCATCCAG
rs469262	ACGTTGGATGTTCTGAGCTGAACGAAGCAG	ACGTTGGATGGGTAGGGATCCTTGTGATGC
rs708323	ACGTTGGATGCACATACTATACAGGTACC	ACGTTGGATGGAGGGAGAAGATGTTGAA
rs469089	ACGTTGGATGTTGGAAAGTACCAACCTCAGC	ACGTTGGATGAATGGAAGGAAGGATCAGCC
rs469396	ACGTTGGATGAGTGAACCTCAATGAGGGAAAC	ACGTTGGATGTCTCACACCAACTGATCCTTC
rs468723	ACGTTGGATGTGGATCTGCTGTTGGG	ACGTTGGATGTATTGGCATCGCGTATCAGG
rs467604	ACGTTGGATGACTCCTGCCATTAAACTCTC	ACGTTGGATGCTGGCTTAACCTACAAGGG
rs338874	ACGTTGGATGCCCAACACAGCCACTGGG	ACGTTGGATGAAGGGCCTGGCCCCACCCAA
rs338875	ACGTTGGATGTGCTGCTTCGCGTGTG	ACGTTGGATGACACTGGATATGTCAGGGTC
rs1385803	ACGTTGGATGTCACCAACCATTCCAGAAGTG	ACGTTGGATGACCTTCCATTGCTGTGGC
rs1385804	ACGTTGGATGTCACCAACCATTCCAGAAGTG	ACGTTGGATGACCTTCCATTGCTGTGGC
rs338876	ACGTTGGATGTTAGGGCTGGGTGGAGGAAG	ACGTTGGATGTCCAACCTCCAGTGACAGAG
rs189803	ACGTTGGATGCCCTCCAGTTCTCTCTTCTG	ACGTTGGATGATCCTGGATTAGCCAGATGG
rs452215	ACGTTGGATGTAGCTCTATTCTCCACCC	ACGTTGGATGAGCGAGACTCCGTCTAAAA
rs641170	ACGTTGGATGATAGCTCTATTCTCCACCC	ACGTTGGATGAGCGAGACTCCGTCTAAAA

dbSNP rs#	Forward PCR primer	Reverse PCR primer
rs584398	ACGTTGGATGTTCCCTGTGAGCTATAAAC	ACGTTGGATGCGAGACTCCGTCTAAAAAAA
rs385330	ACGTTGGATGTTGCCCAACTATTGTCCTG	ACGTTGGATGGGTTCCCAGACAGTGTGG
rs429538	ACGTTGGATGTATTATCTGCAGACACCTGG	ACGTTGGATGATCTCATTCCCACCCCTTC
rs371229	ACGTTGGATGTATTATCTGCAGACACCTGG	ACGTTGGATGATCTCATTCCCACCCCTTC
rs460874	ACGTTGGATGGTCCCTGCGGCTAAAAATTCC	ACGTTGGATGGGCAGGTCAACTAGAAAAC
rs646121	ACGTTGGATGGGCAGGTCAACTAGAAAAC	ACGTTGGATGGTCCCTGCGGCTAAAAATTCC
rs468262	ACGTTGGATGCCAGGTTCGAAAGTTAGG	ACGTTGGATGTGGGTTGGTCATCGCGTAAC
rs467863	ACGTTGGATGTTCGAAACCTGGCTGATGG	ACGTTGGATGTGCCACTGTCAGAAGACAAG
rs191434	ACGTTGGATGCCAGCTGAAACACTAGACAG	ACGTTGGATGAGCTGAAGAGGTCTTCTCC
rs2054782	ACGTTGGATGAAAAAAGCAGGCCAGACC	ACGTTGGATGTCTGACTCTCATCGCAGAC
rs468499	ACGTTGGATGCTCCAGGAGGGACACTACGT	ACGTTGGATGTGGCCAGCTCTCCTCGATG
rs180287	ACGTTGGATGTTCTGCAGAATTACCTAT	ACGTTGGATGGAAAAAGAAAAAAATCAG
rs338877	ACGTTGGATGCCGTGGATGGAAATTACATT	ACGTTGGATGTTCTTGGATCAATGTTGCC
rs650665	ACGTTGGATGCCCATCTTACTCTATGATCTC	ACGTTGGATGAAAGTGCTGGATTATAGGC
rs193419	ACGTTGGATGCCAAATCCAAGACACAGGG	ACGTTGGATGATGTTTCACTACCCCCAGTG
rs180288	ACGTTGGATGTGACCTGGTAGCTTAGAG	ACGTTGGATGTTAGGAGGTCAAGAGGG
rs186834	ACGTTGGATGTAAGCTACCAGGTACACAC	ACGTTGGATGAGTTGATAGGAGAGTCAGGC
rs189266	ACGTTGGATGTAAGCTACCAGGTACACAC	ACGTTGGATGAGTTGATAGGAGAGTCAGGC
rs189267	ACGTTGGATGCCCTATTGTGCCCTGTTGT	ACGTTGGATGCTCTGCCACTCTCCAT
rs170937	ACGTTGGATGCCATCAACTGTTGATGGCG	ACGTTGGATGTTCTCATTGTGCCCTGTTG
rs463263	ACGTTGGATGTACTGGACCCCTTGCACAG	ACGTTGGATGTGCCATGCTCATGTGTTGG
rs463262	ACGTTGGATGCCATGCTCATGTGTTGG	ACGTTGGATGACCCCTTGACAGATGCTG
rs460454	ACGTTGGATGAAGAAGGACCGTGTAGAGA	ACGTTGGATGACATGAGCATGGCAGGTAC
rs460455	ACGTTGGATGACATGAGCATGGCAGGTAC	ACGTTGGATGAAGAAGGACCGTGTAGAGA
rs460505	ACGTTGGATGACCGTGGACAGCGTCTGA	ACGTTGGATGTGCTCTGAGGGCAGAACAAAG
rs931316	ACGTTGGATGATGCACACACCCATGGTCAG	ACGTTGGATGCCGTTACTCCAGCATTTC
rs463431	ACGTTGGATGTCACCACAGCCATGGGA	ACGTTGGATGTTGAAACTACAATGTGGG
rs461542	ACGTTGGATGTGATGAAGGCCAAGAATGCT	ACGTTGGATGTGTCAGAACGTCAAGTG
rs463557	ACGTTGGATGTGATGAAGGCCAAGAATGCT	ACGTTGGATGTGTCAGAACGTCAAGTG
rs191453	ACGTTGGATGCATCCAACAGCTCTGTC	ACGTTGGATGACCCATCTGAGCGCATCAG
rs2271212	ACGTTGGATGAGCTCCCCGGAGGCAACGA	ACGTTGGATGTGCAAGGTCTCGGCCAAAGAC
rs462009	ACGTTGGATGCAGGCTCCCTCGTTGCC	ACGTTGGATGTTGGTGTCCCACGTGGTGT
rs2271211	ACGTTGGATGTCGTACCCCTGCTCTGGACG	ACGTTGGATGACTGACGCCAGGGCCGCTT
rs396474	ACGTTGGATGTTGGAGTTGGAGATGAG	ACGTTGGATGTTCTCAGATCCCAGTCAG
rs428901	ACGTTGGATGTCAGTGACAGAGCGAGACTC	ACGTTGGATGGGCTCGATAATGTAGCCAT
rs452300	ACGTTGGATGAGCACAAGCTGAAGAGGTCT	ACGTTGGATGAGGAGAGAAGTGCACAGATC
rs670256	ACGTTGGATGTAGCTCTATTCTCCACCC	ACGTTGGATGAGCGAGACTCCGTCTAAAA

TABLE 12

dbSNP rs#	Extend Primer	Term Mix
rs2278221	CAAACGCTGAGGAGAACCC	ACT
rs1650358	AAGAGACAAAGGCCGGC	ACT
rs1643818	TACAGGTGTGAACCACCGC	ACT
rs3733916	AGGCTGTAGTGTGACAGAC	ACG

dbSNP rs#	Extend Primer	Term Mix
rs1624933	GTCTCAAACCTCTGACCTCA	ACT
rs1624857	AGACCAGCCTGGCCAACAT	ACT
rs1624832	GGCCAACATGGTAAACCC	ACG
rs1624829	TGGCCAACATGGTAAACCC	ACT
rs2161171	TGGAATAAGAGCCCTGCAGTGG	ACT
rs1530499	CCCCTGCCCGAGCCACAGGAA	ACT
rs888764	ATGTTGATTGGCTATTTGTCA	ACG
rs873987	AAAACCTAAAAGAATCCACGGTA	ACG
rs4078699	GACACATGATTAACAGCAAACAAT	ACT
rs870311	AAGGGCGTGACGGCCCC	ACT
rs1643817	GAAAGGGGAGAAAAGATTATCCC	CGT
rs1643816	AGGACCAGGAGTTCCCATT	ACT
rs1650355	GAATCAATGAAGAAGAGAGCTT	ACT
rs888763	GGTCAGGAGGCAGAGGG	ACT
rs1862212	GGGGTGAAAGGGAGCAGGG	CGT
rs1110514	CAGGGCCCAGGTGAGGAA	CGT
rs3797600	CTTTGTTGTTAACCAAACCC	ACG
rs3797602	GCTGACAGCTCCGGACATG	ACT
rs3797603	TGTCATTCTCCTTGTGAACCCCTC	ACT
rs3776819	CCATTCATTGCACCTGCATG	ACT
rs252076	CAAAGTGCTGGATTGCAGG	ACG
rs252075	GAGCATTGCAGGCATGCCCTCT	ACT
rs252074	CTGGGTGGCTGCTGGGC	ACG
rs252068	GGAAGGGTTAACGAAAGGAG	ACT
rs252069	TGAGCACCTACTATGGGCTAG	ACT
rs194040	ATTCCATATCTTCAAAGTGATTCA	ACG
rs252070	CCTGGGCTTCCCTCCC	ACG
rs3797606	AGCCCTTGGCCTCTCTCC	ACT
rs171667	CGCCTTTGCTTATGAAAGA	ACG
rs187539	AACCTTCAGGAAAGTCCCCT	ACT
rs3836834	TCAAAATATCAAACATACCATGAAA	ACG
rs252071	ACCTGAGACACAGGGACT	ACT
rs252072	GCTGGGTACACTCGCGGA	ACG
rs252073	GCCGAGAAAAGTCAGGGATTCT	ACT
rs379589	CGGGAGTACTGAGCACCCAGG	CGT
rs2052472	CCCCACTGTACTATCTCCAC	ACT
rs2052471	GTCTCCTTGGCTGCCAAG	ACT
rs2052470	TGCCAAGGCCCTGTCCTC	ACT
rs2052469	CGCGGGGAAGTACTCGGC	ACT
rs3797608	GTCCTCCTGTTCTGAGGCC	ACT
rs3797609	GGCAGAGCGGATGGCCTG	ACG
rs3822601	GTGAGGCCTGAGATGAGAAC	ACG
rs153131	TCCCCATACCTGTGCTC	ACG
rs751546	CCGTTCTACAGCGGTTAAGA	ACT

dbSNP rs#	Extend Primer	Term Mix
rs2279979	GGCCACCAGACAGATGTAAG	ACT
rs252060	CGTGTTCGGCAGAGGTGA	ACT
rs3797610	CTTCTCCCTGGGTATGTGTT	ACT
rs194039	CCACCGTCCCAGGACATTTTTTT	ACT
rs168773	ACTGGAGATATCACGGGAGC	CGT
rs252061	CCAGCTGGTCACAGGGCTCCC	ACG
rs187537	TCATGCTAAGTCAAATAAGCCA	ACT
rs252062	GCCACCACCGTCCACAGA	CGT
rs2431255	CTGTATATTCACCGCAATTAAAAA	ACT
rs3797612	GGCATTCATCGTCAGGGCAA	ACG
rs3797613	CCGCCGCCGGTCTCCCA	ACG
rs614114	GAACGTTCTCTCACTTTGCC	ACT
rs252063	CTCCCGTCCTCTGAGCCTT	ACT
rs252064	GAAAACTAAGGCTCAGAGGAC	ACT
rs252065	TGGAAAAGGCGAGGCCTGGAGT	ACG
rs450502	GGAAAAGGCGAGGCCTGGAGTT	ACT
rs439252	GCCTCCCAAAGTGTGGGATTA	ACG
rs252066	TAGCCCTCTGGAGGCCAG	ACG
rs457957	GGGCCCTCCTAAAGCTC	ACT
rs3797614	TGGCCCTCGCTCTAACGCA	ACG
rs423552	CTCGATGTTGTAGTCATCGTC	ACG
rs398829	TGGCGTGCTCCTCTAGGA	ACG
rs416646	CTCAGCAGGTCTGATCCATC	ACT
rs187450	GGGCAGACTCCCCAGGAT	ACT
rs337807	GCAGGCCACTCGGTGGAC	ACT
rs337806	CCACCCCAGGGTAGCCC	ACT
rs1396438	GGCAGGCAGGTGGCCTG	ACT
rs1396437	CGGCAGAAGCAGCCTAAAGA	ACG
rs2411811	ACATAATTCCAAATTCACCCC	CGT
rs2898813	GGTCCTGGGTGGAGGGAT	ACT
rs189256	AGCAAGGCTCTATTGGGGA	ACT
rs173072	CTTGCTCACATCGTGGCCAAA	ACT
rs337805	GGAGCAGGAAAATTACATGACT	ACG
rs191415	GAGTCCAACGTTGTCCACAT	ACT
rs180045	ACTTGTTCCTACAATTCTCATT	ACG
rs189255	GACTCAGTCCCAGGTTTCT	ACT
rs652766	AGAAAACCTGGGACTGAGTCT	ACT
rs466750	TCCTTAAATGCCTGGTGGCAAT	ACT
rs442406	TTCTGGCTGTTGGGTTGAAC	ACT
rs662407	AAACATCTGAAATTAAAGCACC	ACT
rs592971	AGCATTCTTGTACTGCTTTCA	ACT
rs457187	GGAGTATCTGTTCTGTGG	ACT
rs459490	TTGAACATAGGAATAACCCGC	ACT
rs459668	GTCTTCTTGTGTTGGAGA	ACG

dbSNP rs#	Extend Primer	Term Mix
rs462646	ATTATTCGACGGAGATTATTGAC	ACT
rs458272	ATTATTTCTGTCTGGTGTGG	ACT
rs463455	CCTCCTCATTGTCAATTCTTTTC	ACT
rs675880	CTTTCATGACATTGACACAAC	ACT
rs810617	CCACAGCCTCCGCCTCCC	ACT
rs464156	GGGTTTCCAGGTTAAATGGC	ACT
rs458083	CTCCTGCTCTGCCTATCCTT	ACT
rs467333	ATTTCTTCCCCTCCTGCTCT	ACT
rs465381	GCCTCCCACAGTCCCTTGTT	ACT
rs466363	GTGATGGCTCTGCACCAGA	ACG
rs2457099	AGCGTGTGCCAGCCTCTCC	ACT
rs463901	GACACAATTCAAGAGCGACTTAC	ACT
rs465621	AAGTGCTAGAAGAAAATGTAGC	ACT
rs463724	CCTTGCGCCATCCCCCTAG	CGT
rs465242	TGTGCCCATCCCCCCCCTT	ACT
rs467419	AACGTAAGTCCTAATGACCGCCC	ACG
rs456135	CCCTCTCCTTTCTGGGCA	ACT
rs464536	GCTTTGGTCCTCCTGAGCC	ACG
rs461898	CACAGAGCGACTCTCTTGGTT	ACT
rs389558	CTGACAGTTCTCTAAACTCCA	ACG
rs466752	TTCCTTTCTCCCTGAAGCA	ACG
rs455655	CACCCCATATGGCTCATGGG	ACT
rs463435	GGGAAGGAGGTACTTAGCAG	ACG
rs2174971	GTGCCACTCTCCAGCGGCC	ACG
rs1979979	TTGCCGGCCCCCACCTC	ACG
rs411804	GAAAATCCCTGTCACCAGTC	ACG
rs1623885	CTTGGCTGCAGCACCCCA	ACT
rs1643811	GCCCTGGTACTTCAGCTCCCT	ACG
rs434430	GTGTGCATGTGTGTGCCTG	CGT
rs187538	TAAACGGGCCAAAAACGCCAT	ACT
rs252067	GCGCCTACGGATGTCA	ACT
rs459319	CATGTTAACAGAGAGAACGGTC	ACG
rs467289	TCACTGAGAAATATTTGCTCCC	ACT
rs462644	GGTGAGCCCTGGCTGTG	ACG
rs458752	TAAAGCGCTTACAATCAACA	ACT
rs708320	TGGTGAATCCTGCTCTACTAAA	CGT
rs457954	CACCGTTCTTATAATGCAGCC	ACT
rs2411810	GGGGACGTTACCTTTTCA	ACG
rs3084687	ATTTATATATGTGTGTACACAT	ACT
rs69638	CCCATTGGCTGTCTGGAA	ACT
rs455452	CCTCAACCCCAAGATGCCCTC	ACG
rs464850	ACTCCTGCCTGAGTGTCTC	ACT
rs431472	GTGAAGCGGAAGGAGACTC	ACG
rs2411809	CTGCACACCCCTGACAG	ACG

dbSNP rs#	Extend Primer	Term Mix
rs2457094	TGGCTGGCACCACTGCACTGC	ACT
rs2457095	TGGCTCATGCTTCTAATCCCA	ACT
rs2261740	CTGCACTCCAGCCTGGGC	ACT
rs1109180	ACATCAGTGACAGTGTAAATGGTA	ACG
rs1109179	TATGCAGACCCCCCTCCCC	ACT
rs1109178	AACAACAGCAACAGAAAATGAAG	ACT
rs456909	CGATTCCCACCGCGTGTCTG	ACG
rs469124	CCTGGCTCCATTGGTGTGAA	ACT
rs468039	CCTTCACACCAATGGAGCCAG	ACT
rs467017	CTGCCACTAAACAGATGAGAA	ACT
rs469290	ATTTCTGGGCCCAAAGTCCA	ACT
rs469090	CCAATTGTTCCAGCCACTCCC	ACT
rs469568	TGATATTGCTTGCTTGGTCTTAG	ACT
rs468386	GGTCAAGAACATTCAAGAGGCAGC	ACT
rs469349	GTGCAGTGGCACGATCCTA	ACT
rs469099	GCAGGTGGAACCGCAGAC	ACT
rs456868	GGAGCTGCGGTGACTCCC	ACT
rs465389	CCCTGGCACTCGCAGACC	ACT
rs463892	AGCTCCCCCGCACCAAC	ACT
rs468548	AAGGGAACATGCAAGCAAAGACTC	ACT
rs654612	TGCAAGCAAAGACTCGAATGA	ACT
rs468542	TCACTCACTTGATTCCCTGCCATC	ACT
rs469262	CACTGTGGGATTCCAGCAGA	ACT
rs708323	TATACAGGTACCCATTAAAGT	ACT
rs469089	CCTCGGCCTCCCCAGCT	ACT
rs469396	AATGAGGGAACCTGCAGTTAAGA	ACT
rs468723	CAGACCCATGCCTTGCC	ACT
rs467604	GAGTTCCCTCCTTTACAA	ACT
rs338874	CACAGCCACTGGGGAGTAG	ACT
rs338875	TTGCTCGCGTGTGCCAGCAAAT	ACG
rs1385803	AAGTGGAATTCTCATGGCAGAT	ACT
rs1385804	CATTCCAGAAGTGGAATTCTCATG	ACT
rs338876	AGGAAGGTGCTCCGGCCT	ACG
rs189803	TGCTTCCCCCTCCCCCT	CGT
rs452215	TCTATTCTCCACCCCCATCTT	ACT
rs641170	CTATTCTCCACCCCCATCT	ACT
rs584398	CTCTTATATAGCTCTATTCTCC	CGT
rs385330	AGGTGTCTGCAGATAATACATT	ACG
rs429538	CCTGGGGCACAGGACAATA	ACT
rs371229	GACAATAGTTGGGGCAAGAC	ACT
rs460874	ACAAAACATCCTTCAAAAATACA	CGT
rs646121	GTTTTGTTCTCTGAAAGTGTCT	CGT
rs468262	CACCCAACACTATTGCTCCC	ACG
rs467863	GCTGATGGGAGGCCAATGT	ACT

dbSNP rs#	Extend Primer	Term Mix
rs191434	GTCCAGAGATCCTGCTCACT	CGT
rs2054782	CCCCCTCCATCACCTCCC	ACG
rs468499	GTGAGCCAGCAATTCTCCTA	ACT
rs180287	CAATGATCAGAACTCAGAGGTTT	ACT
rs338877	AGAGATAAATTCCAGTGTGAG	CGT
rs650665	AGACATCCCAGGCCGGGC	ACT
rs193419	CCAAAGACACAGGGAGTAGATTA	ACT
rs180288	GAGAATATTCTTGTGGGCTTAAT	ACT
rs186834	CCAGGTACACACACACTC	ACG
rs189266	CACACACTCCCTCTCACTGT	ACT
rs189267	TTCTGTGCATCTTGACGCCATC	CGT
rs170937	GATGGCGTCAAAGATGCACA	ACT
rs463263	CCCCTTGACAGATGCTG	ACT
rs463262	GGGGAGCAGCCAGTTCCTA	ACT
rs460454	AGAGGCTGGGGACAGAGAA	ACT
rs460455	GGTACCCACCAGTCTCCTTCT	ACT
rs460505	CAGCGTCTCTGACACGGTC	ACG
rs931316	GGTCAGAGCAGACACATCCACAT	ACG
rs463431	CCCATGGGGAGCACCAAG	ACT
rs461542	TGGGAGCTCCGGGATATTGCC	ACG
rs463557	GCTCCCAGGATATTGCCCA	ACT
rs191453	CTGGGCTGGGGCCCTGC	ACT
rs2271212	CGAGGGAGGAGCCTGGCAG	ACG
rs462009	CTCCTCGTTGCCTCCGGG	ACT
rs2271211	GACGTAGCTGCCGACACCA	ACG
rs396474	CTGGTGGCCCATCTATCCTGG	ACT
rs428901	GAGCGAGACTCCGTCTCAA	CGT
rs452300	CTGAAGAGGTCTTCTCCTTCC	CGT
rs670256	TTCTTCCACCCCCATCTTG	ACT

Genetic Analysis

[0237] Allelotyping results from the discovery cohort are shown for cases and controls in Table 13. The allele frequency for the A2 allele is noted in the fifth and sixth columns for osteoarthritis case pools and control pools, respectively, where “AF” is allele frequency. The allele frequency for the A1 allele can be easily calculated by subtracting the A2 allele frequency from 1 (A1 AF = 1-A2 AF). For example, the SNP rs2278221 has the following case and control allele frequencies: case A1 (C) = 0.36; case A2 (T) = 0.64; control A1 (C) = 0.37; and control A2 (T) = 0.63, where the nucleotide is provided in parenthesis. Some SNPs are labeled “untyped” because of failed assays.

TABLE 13

dbSNP rs#	Position in SEQ ID NO: 1	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs2278221	210	178695460	C/T	0.64	0.63	0.770
rs1650358	3608	178698858	C/G			
rs1643818	3609	178698859	C/G			
rs3733916	4318	178699568	C/T			
rs1624933	5593	178700843	A/G	0.69	0.71	0.255
rs1624857	5629	178700879	C/T	0.79	0.81	0.574
rs1624832	5639	178700889	A/G	0.41	0.44	0.203
rs1624829	5640	178700890	C/T	0.89	0.93	0.044
rs2161171	8943	178704193	A/C			
rs1530499	17968	178713218	A/G	0.39	0.39	0.861
rs888764	19887	178715137	A/G			
rs873987	21034	178716284	A/G			
rs4078699	21085	178716335	C/T	0.56	0.54	0.374
rs870311	21596	178716846	A/G	0.51	0.50	0.590
rs1643817	23379	178718629	A/C	0.27	NA	NA
rs1643816	23432	178718682	A/C			
rs1650355	24007	178719257	A/C			
rs888763	26121	178721371	A/G	0.40	0.42	0.390
rs1862212	26273	178721523	A/T	0.55	0.54	0.753
rs1110514	26755	178722005	A/T	0.29	0.28	0.572
rs3797600	27411	178722661	C/T	0.56	0.57	0.738
rs3797602	27710	178722960	G/T	0.65	0.64	0.564
rs3797603	27842	178723092	C/T			
rs3776819	28379	178723629	C/T	0.46	0.46	0.850
rs252076	29603	178724853	C/T	0.46	0.48	0.519
rs252075	31232	178726482	C/G	0.35	0.36	0.859
rs252074	31504	178726754	A/G	0.35	0.34	0.816
rs252068	32583	178727833	C/G	0.47	0.48	0.656
rs252069	32794	178728044	A/G	0.28	0.27	0.626
rs194040	32840	178728090	C/T	0.31	0.32	0.665
rs252070	33044	178728294	C/T	0.58	0.57	0.573
rs3797606	33150	178728400	A/C	0.88	0.88	0.684
rs171667	33218	178728468	A/G	0.48	0.51	0.166
rs187539	33513	178728763	C/T	0.33	0.34	0.652
rs3836834	33959	178729209	- /TATCA AACTAC CATGAA A			
rs252071	34486	178729736	A/G	0.30	0.31	0.666
rs252072	36289	178731539	C/T	0.49	0.50	0.677
rs252073	36570	178731820	C/T			
rs379589	38247	178733497	A/T	0.59	0.63	0.096
rs2052472	38477	178733727	A/C	0.05	0.06	0.508
rs2052471	38518	178733768	C/T	0.89	0.88	0.459
rs2052470	38529	178733779	C/T	0.83	0.80	0.125
rs2052469	38667	178733917	A/G	0.83	0.80	0.172
rs3797608	39781	178735031	C/T	0.06	0.07	0.578
rs3797609	39856	178735106	C/T	0.05	0.05	0.812
rs3822601	39927	178735177	C/T	0.08	0.08	0.802
rs153131	40506	178735756	A/G	0.76	0.77	0.944
rs751546	41869	178737119	C/G	0.93	0.92	0.585
rs2279979	42452	178737702	C/T	0.93	0.92	0.436
rs252060	44788	178740038	C/T	0.81	0.82	0.760
rs3797610	46059	178741309	A/C	0.17	0.17	0.858

dbSNP rs#	Position in SEQ ID NO: 1	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs194039	46846	178742096	A/G	0.41	0.47	0.035
rs168773	47712	178742962	A/T	0.35	0.38	0.266
rs252061	48796	178744046	C/T	0.21	0.19	0.508
rs187537	49441	178744691	C/G			
rs252062	49602	178744852	A/T	0.95	0.95	0.960
rs2431255	49723	178744973	A/C	0.24	0.19	0.034
rs3797612	50050	178745300	C/T	0.38	0.43	0.036
rs3797613	50171	178745421	C/T	0.21	0.21	0.941
rs614114	50477	178745727	C/T	0.50	0.53	0.387
rs252063	50818	178746068	C/T	0.57	0.55	0.313
rs252064	50833	178746083	C/T	0.52	0.52	0.806
rs252065	50881	178746131	A/G	0.22	0.22	0.857
rs450502	50882	178746132	A/G			
rs439252	51386	178746636	C/T			
rs252066	51534	178746784	C/T	0.19	0.18	0.618
rs457957	52317	178747567	A/G	0.67	0.70	0.172
rs3797614	52368	178747618	C/T			
rs423552	52970	178748220	A/G	0.90	0.92	0.215
rs398829	53023	178748273	A/G			
rs416646	53356	178748606	A/G	0.56	0.57	0.650
rs187450	53882	178749132	G/T			
rs337807	54553	178749803	C/T	0.55	0.59	0.208
rs337806	55475	178750725	A/C	0.11	0.10	0.925
rs1396438	55530	178750780	A/G	0.56	0.54	0.494
rs1396437	55691	178750941	C/T			
rs2411811	55848	178751098	A/C			
rs2898813	55879	178751129	C/G			
rs189256	56316	178751566	A/G	0.19	0.19	0.988
rs173072	56911	178752161	A/C			
rs337805	57320	178752570	A/G	0.25	0.24	0.657
rs191415	57391	178752641	C/T			
rs180045	57437	178752687	C/T	0.51	0.47	0.211
rs189255	57478	178752728	C/G	0.15	0.12	0.273
rs652766	57500	178752750	C/T	0.57	0.61	0.213
rs466750	59111	178754361	G/T	0.35	0.33	0.493
rs442406	59333	178754583	A/G	0.57	0.59	0.420
rs662407	59715	178754965	A/G	0.31	0.27	0.102
rs592971	59804	178755054	A/G			
rs457187	59851	178755101	A/G	0.23	0.24	0.842
rs459490	59929	178755179	C/T	0.21	0.20	0.604
rs459668	60052	178755302	C/T	0.20	0.19	0.648
rs462646	60240	178755490	C/T	0.43	0.43	0.905
rs458272	60359	178755609	G/T	0.22	0.20	0.523
rs463455	60381	178755631	A/G	0.25	0.24	0.644
rs675880	60456	178755706	C/T	0.63	0.65	0.591
rs810617	60724	178755974	C/G			
rs464156	60875	178756125	C/T	0.34	0.34	0.892
rs458083	60968	178756218	A/G	0.80	0.82	0.499
rs467333	60978	178756228	C/G	0.11	0.12	0.369
rs465381	60998	178756248	C/T			
rs466363	61557	178756807	C/T	0.31	0.34	0.358
rs2457099	62091	178757341	C/T	0.44	0.44	0.956
rs463901	62645	178757895	C/T	0.43	0.45	0.395
rs465621	62943	178758193	A/C	0.62	0.63	0.534
rs463724	63131	178758381	A/T	0.09	0.08	0.523
rs465242	63145	178758395	G/T			
rs467419	63406	178758656	A/G	0.65	0.66	0.647

dbSNP rs#	Position in SEQ ID NO: 1	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs456135	63427	178758677	C/G	0.79	0.80	0.686
rs464536	63554	178758804	C/T	0.36	0.34	0.296
rs461898	63661	178758911	A/G	0.30	0.32	0.411
rs389558	64093	178759343	A/G	0.24	0.26	0.325
rs466752	64153	178759403	C/T	0.35	0.37	0.446
rs455655	64409	178759659	C/G	0.87	0.89	0.536
rs463435	64544	178759794	C/T	0.68	0.66	0.428
rs2174971	65257	178760507	C/T	0.52	0.51	0.695
rs1979979	65626	178760876	A/G	0.07	0.06	0.692
rs411804	65739	178760989	A/G	0.78	0.78	0.976
rs1623885	66392	178761642	C/T	0.82	0.80	0.492
rs1643811	66720	178761970	C/T	0.24	0.24	0.924
rs434430	69177	178764427	A/T			
rs187538	69336	178764586	G/T			
rs252067	69636	178764886	A/G	0.21	0.23	0.606
rs459319	69823	178765073	A/G	0.19	0.20	0.640
rs467289	69928	178765178	C/T	0.26	0.26	0.988
rs462644	70547	178765797	C/T	0.59	0.58	0.914
rs458752	70633	178765883	C/T	0.18	0.20	0.513
rs708320	71805	178767055	A/C			
rs457954	72181	178767431	C/G	0.71	0.73	0.327
rs2411810	72200	178767450	C/T	0.28	0.26	0.252
rs3084687	72474	178767724	-/AT	0.13	0.12	0.884
rs69638	72567	178767817	C/G	0.54	0.52	0.449
rs455452	72973	178768223	A/G	0.59	0.60	0.733
rs464850	73468	178768718	A/G	0.11	0.09	0.249
rs431472	73889	178769139	A/G	0.33	0.34	0.713
rs2411809	75730	178770980	C/T			
rs2457094	75970	178771220	A/G	0.71	0.73	0.383
rs2457095	76114	178771364	A/G	0.74	0.76	0.551
rs2261740	76342	178771592	C/T	0.35	0.36	0.702
rs1109180	76449	178771699	A/G			
rs1109179	76465	178771715	C/T			
rs1109178	76791	178772041	A/C	0.46	0.45	0.820
rs456909	78042	178773292	A/G	0.55	0.53	0.444
rs469124	80758	178776008	A/G			
rs468039	80778	178776028	C/T			
rs467017	81356	178776606	A/C	0.33	0.32	0.665
rs469290	81576	178776826	A/G	0.57	0.57	0.871
rs469090	81689	178776939	C/T	0.82	0.83	0.387
rs469568	81759	178777009	G/T	0.38	0.38	0.888
rs468386	81950	178777200	C/G			
rs469349	82562	178777812	A/C			
rs469099	83591	178778841	C/T	0.66	0.63	0.264
rs456868	83700	178778950	A/G			
rs465389	83821	178779071	C/G			
rs463892	83842	178779092	C/G			
rs468548	83923	178779173	G/T			
rs654612	83929	178779179	A/C			
rs468542	84021	178779271	C/G			
rs469262	84175	178779425	C/T	0.45	0.47	0.405
rs708323	84417	178779667	A/G	0.73	0.69	0.138
rs469089	84747	178779997	C/G			
rs469396	85746	178780996	C/G	0.38	0.37	0.817
rs468723	86129	178781379	C/T	0.37	0.38	0.754
rs467604	86335	178781585	A/G	0.34	0.32	0.504
rs338874	87315	178782565	C/G	0.43	0.44	0.879

dbSNP rs#	Position in SEQ ID NO: 1	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs338875	87648	178782898	A/G	0.48	0.50	0.289
rs1385803	87764	178783014	A/C			
rs1385804	87770	178783020	C/G			
rs338876	88221	178783471	C/T	0.39	0.39	0.889
rs189803	90474	178785724	A/C			
rs452215	91148	178786398	G/T			
rs641170	91150	178786400	G/T			
rs584398	91160	178786410	G/T			
rs385330	91733	178786983	C/T			
rs429538	91772	178787022	A/C			
rs371229	91785	178787035	C/T			
rs460874	93140	178788390	A/T	0.74	0.71	0.351
rs646121	93148	178788398	A/T	0.93	0.94	0.687
rs468262	96080	178791330	A/G			
rs467863	96157	178791407	C/G			
rs191434	96313	178791563	A/C			
rs2054782	96759	178792009	C/T	0.44	0.42	0.353
rs468499	97026	178792276	A/C			
rs180287	97320	178792570	C/G			
rs338877	97732	178792982	A/T	0.04	0.04	0.863
rs650665	98713	178793963	C/G			
rs193419	99707	178794957	A/C			
rs180288	99959	178795209	C/G			
rs186834	100009	178795259	A/G			
rs189266	100020	178795270	C/G			
rs189267	100065	178795315	A/C			
rs170937	100086	178795336	C/G			
rs463263	101270	178796520	C/G			
rs463262	101276	178796526	G/T			
rs460454	101371	178796621	C/T			
rs460455	101376	178796626	C/G			
rs460505	101439	178796689	C/T			
rs931316	101820	178797070	C/T			
rs463431	102392	178797642	C/G			
rs461542	102602	178797852	A/G			
rs463557	102604	178797854	A/C			
rs191453	102896	178798146	C/T	0.11	0.14	0.123
rs2271212	189104	178884354	C/T	0.65	0.57	0.003
rs462009	189134	178884384	C/T			
rs2271211	189205	178884455	A/G			
rs396474	Not mapped	Not mapped	A/C			
rs428901	Not mapped	Not mapped	A/T	0.64	0.72	0.015
rs452300	Not mapped	Not mapped	G/T			
rs670256	Not mapped	Not mapped	G/T			

[0238] The *ADAMTS2* proximal SNPs were also allelotyped in the replication cohorts using the methods described herein and the primers provided in Tables 11 and 12. The replication allelotyping results for replication cohort #1 and replication cohort #2 are provided in Tables 14 and 15, respectively.

TABLE 14

dbSNP rs#	Position in SEQ ID NO: 1	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p-Value
rs2278221	210	178695460	C/T	0.64	0.62	0.624
rs1650358	3608	178698858	C/G			
rs1643818	3609	178698859	C/G			
rs3733916	4318	178699568	C/T			
rs1624933	5593	178700843	A/G	0.65	0.69	0.322
rs1624857	5629	178700879	C/T	0.81	untyped	NA
rs1624832	5639	178700889	A/G	0.38	0.42	0.265
rs1624829	5640	178700890	C/T	0.87	untyped	NA
rs2161171	8943	178704193	A/C			
rs1530499	17968	178713218	A/G	0.39	0.40	0.765
rs888764	19887	178715137	A/G			
rs873987	21034	178716284	A/G			
rs4078699	21085	178716335	C/T	0.55	0.54	0.733
rs870311	21596	178716846	A/G	0.50	0.50	0.828
rs1643817	23379	178718629	A/C	0.27	untyped	
rs1643816	23432	178718682	A/C			
rs1650355	24007	178719257	A/C			
rs888763	26121	178721371	A/G	0.40	0.40	0.816
rs1862212	26273	178721523	A/T	0.55	0.55	0.936
rs1110514	26755	178722005	A/T	0.29	0.29	0.997
rs3797600	27411	178722661	C/T	0.57	0.58	0.604
rs3797602	27710	178722960	G/T	0.64	0.63	0.879
rs3797603	27842	178723092	C/T			
rs3776819	28379	178723629	C/T	0.47	0.46	0.889
rs252076	29603	178724853	C/T	0.46	0.49	0.410
rs252075	31232	178726482	C/G	0.35	0.37	0.572
rs252074	31504	178726754	A/G	0.35	0.35	0.914
rs252068	32583	178727833	C/G	0.48	0.48	0.853
rs252069	32794	178728044	A/G	0.29	0.28	0.765
rs194040	32840	178728090	C/T	0.31	0.33	0.450
rs252070	33044	178728294	C/T	0.57	0.58	0.609
rs3797606	33150	178728400	A/C	0.87	0.91	0.119
rs171667	33218	178728468	A/G	0.45	0.50	0.125
rs187539	33513	178728763	C/T	0.33	0.34	0.709
rs3836834	33959	178729209	- TATCA AACTAC CATGAA A			
rs252071	34486	178729736	A/G	0.30	0.32	0.566
rs252072	36289	178731539	C/T	0.48	0.51	0.400
rs252073	36570	178731820	C/T			
rs379589	38247	178733497	A/T	0.59	0.65	0.035
rs2052472	38477	178733727	A/C	0.04	0.06	0.493
rs2052471	38518	178733768	C/T	0.87	0.88	0.697
rs2052470	38529	178733779	C/T	0.84	0.78	0.036
rs2052469	38667	178733917	A/G	0.84	0.79	0.086
rs3797608	39781	178735031	C/T	0.06	0.07	0.530
rs3797609	39856	178735106	C/T	0.04	0.05	0.841
rs3822601	39927	178735177	C/T	0.08	0.08	0.904
rs153131	40506	178735756	A/G	0.77	0.77	0.964
rs751546	41869	178737119	C/G	0.94	0.92	0.265
rs2279979	42452	178737702	C/T	0.94	0.92	0.238
rs252060	44788	178740038	C/T	0.82	0.80	0.553
rs3797610	46059	178741309	A/C	0.16	0.18	0.459

dbSNP rs#	Position in SEQ ID NO: 1	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs194039	46846	178742096	A/G	0.43	0.45	0.589
rs168773	47712	178742962	A/T	0.34	0.35	0.845
rs252061	48796	178744046	C/T	0.23	0.22	0.884
rs187537	49441	178744691	C/G			
rs252062	49602	178744852	A/T	0.98	0.96	0.310
rs2431255	49723	178744973	A/C	0.24	0.19	0.108
rs3797612	50050	178745300	C/T	0.42	0.46	0.254
rs3797613	50171	178745421	C/T	0.19	0.21	0.576
rs614114	50477	178745727	C/T	0.52	0.54	0.717
rs252063	50818	178746068	C/T	0.55	0.57	0.537
rs252064	50833	178746083	C/T	0.52	0.50	0.609
rs252065	50881	178746131	A/G	0.21	0.25	0.234
rs450502	50882	178746132	A/G			
rs439252	51386	178746636	C/T			
rs252066	51534	178746784	C/T	0.20	0.20	0.883
rs457957	52317	178747567	A/G	0.66	0.71	0.162
rs3797614	52368	178747618	C/T			
rs423552	52970	178748220	A/G	0.90	0.92	0.380
rs398829	53023	178748273	A/G			
rs416646	53356	178748606	A/G	0.58	0.59	0.915
rs187450	53882	178749132	G/T			
rs337807	54553	178749803	C/T	0.60	NA	NA
rs337806	55475	178750725	A/C	0.10	0.10	0.997
rs1396438	55530	178750780	A/G	0.52	0.57	0.188
rs1396437	55691	178750941	C/T			
rs2411811	55848	178751098	A/C			
rs2898813	55879	178751129	C/G			
rs189256	56316	178751566	A/G	0.21	0.20	0.852
rs173072	56911	178752161	A/C			
rs337805	57320	178752570	A/G	0.24	0.24	0.950
rs191415	57391	178752641	C/T			
rs180045	57437	178752687	C/T	0.47	0.46	0.918
rs189255	57478	178752728	C/G	0.14	0.13	0.764
rs652766	57500	178752750	C/T	0.59	0.61	0.570
rs466750	59111	178754361	G/T	0.38	0.37	0.606
rs442406	59333	178754583	A/G	0.56	0.57	0.882
rs662407	59715	178754965	A/G	0.32	0.27	0.134
rs592971	59804	178755054	A/G			
rs457187	59851	178755101	A/G	0.23	0.25	0.451
rs459490	59929	178755179	C/T	0.22	0.21	0.671
rs459668	60052	178755302	C/T	0.20	0.19	0.712
rs462646	60240	178755490	C/T	0.42	0.44	0.439
rs458272	60359	178755609	G/T	0.21	0.21	0.755
rs463455	60381	178755631	A/G	0.25	0.25	0.783
rs675880	60456	178755706	C/T	0.62	0.63	0.741
rs810617	60724	178755974	C/G			
rs464156	60875	178756125	C/T	0.32	0.34	0.541
rs458083	60968	178756218	A/G	0.80	0.82	0.499
rs467333	60978	178756228	C/G	0.10	0.13	0.243
rs465381	60998	178756248	C/T			
rs466363	61557	178756807	C/T	0.31	0.34	0.494
rs2457099	62091	178757341	C/T	0.45	0.45	0.997
rs463901	62645	178757895	C/T	0.46	0.46	0.852
rs465621	62943	178758193	A/C	0.64	0.63	0.853
rs463724	63131	178758381	A/T	0.09	0.08	0.737
rs465242	63145	178758395	G/T			
rs467419	63406	178758656	A/G	0.64	0.65	0.694

dbSNP rs#	Position in SEQ ID NO: 1	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs456135	63427	178758677	C/G	0.79	0.76	0.339
rs464536	63554	178758804	C/T	0.36	0.34	0.553
rs461898	63661	178758911	A/G	0.31	0.33	0.727
rs389558	64093	178759343	A/G	0.27	0.28	0.762
rs466752	64153	178759403	C/T	0.34	0.38	0.223
rs455655	64409	178759659	C/G	0.87	untyped	NA
rs463435	64544	178759794	C/T	0.65	0.65	0.973
rs2174971	65257	178760507	C/T	0.49	0.51	0.476
rs1979979	65626	178760876	A/G	0.08	0.07	0.579
rs411804	65739	178760989	A/G	0.77	0.79	0.420
rs1623885	66392	178761642	C/T	0.81	0.78	0.451
rs1643811	66720	178761970	C/T	0.26	0.25	0.715
rs434430	69177	178764427	A/T			
rs187538	69336	178764586	G/T			
rs252067	69636	178764886	A/G	0.22	0.22	0.978
rs459319	69823	178765073	A/G	0.19	0.22	0.245
rs467289	69928	178765178	C/T	0.26	0.29	0.377
rs462644	70547	178765797	C/T	0.58	0.56	0.637
rs458752	70633	178765883	C/T	0.18	0.23	0.129
rs708320	71805	178767055	A/C			
rs457954	72181	178767431	C/G	0.69	0.73	0.143
rs2411810	72200	178767450	C/T	0.28	0.23	0.083
rs3084687	72474	178767724	-/AT	0.12	0.13	0.767
rs69638	72567	178767817	C/G	0.53	0.49	0.157
rs455452	72973	178768223	A/G	0.58	0.61	0.313
rs464850	73468	178768718	A/G	0.13	0.10	0.171
rs431472	73889	178769139	A/G	0.32	0.39	0.048
rs2411809	75730	178770980	C/T			
rs2457094	75970	178771220	A/G	0.70	0.75	0.157
rs2457095	76114	178771364	A/G	0.74	0.75	0.707
rs2261740	76342	178771592	C/T	0.34	untyped	NA
rs1109180	76449	178771699	A/G			
rs1109179	76465	178771715	C/T			
rs1109178	76791	178772041	A/C	0.47	0.48	0.715
rs456909	78042	178773292	A/G	0.56	0.54	0.537
rs469124	80758	178776008	A/G			
rs468039	80778	178776028	C/T			
rs467017	81356	178776606	A/C	0.33	0.31	0.480
rs469290	81576	178776826	A/G	0.63	0.66	0.427
rs469090	81689	178776939	C/T	0.80	0.83	0.300
rs469568	81759	178777009	G/T	0.39	0.43	0.234
rs468386	81950	178777200	C/G			
rs469349	82562	178777812	A/C			
rs469099	83591	178778841	C/T	0.66	0.60	0.066
rs456868	83700	178778950	A/G			
rs465389	83821	178779071	C/G			
rs463892	83842	178779092	C/G			
rs468548	83923	178779173	G/T			
rs654612	83929	178779179	A/C			
rs468542	84021	178779271	C/G			
rs469262	84175	178779425	C/T	0.46	0.50	0.232
rs708323	84417	178779667	A/G	0.72	0.66	0.071
rs469089	84747	178779997	C/G			
rs469396	85746	178780996	C/G	0.37	0.35	0.522
rs468723	86129	178781379	C/T	0.39	0.41	0.495
rs467604	86335	178781585	A/G	0.33	0.30	0.303
rs338874	87315	178782565	C/G	0.44	0.46	0.628

dbSNP rs#	Position in SEQ ID NO: 1	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs338875	87648	178782898	A/G	0.49	0.54	0.106
rs1385803	87764	178783014	A/C			
rs1385804	87770	178783020	C/G			
rs338876	88221	178783471	C/T	0.38	0.36	0.609
rs189803	90474	178785724	A/C			
rs452215	91148	178786398	G/T			
rs641170	91150	178786400	G/T			
rs584398	91160	178786410	G/T			
rs385330	91733	178786983	C/T			
rs429538	91772	178787022	A/C			
rs371229	91785	178787035	C/T			
rs460874	93140	178788390	A/T	0.74	0.69	0.118
rs646121	93148	178788398	A/T	0.93	0.95	0.477
rs468262	96080	178791330	A/G			
rs467863	96157	178791407	C/G			
rs191434	96313	178791563	A/C			
rs2054782	96759	178792009	C/T	0.45	0.42	0.514
rs468499	97026	178792276	A/C			
rs180287	97320	178792570	C/G			
rs338877	97732	178792982	A/T	0.04	0.04	0.781
rs650665	98713	178793963	C/G			
rs193419	99707	178794957	A/C			
rs180288	99959	178795209	C/G			
rs186834	100009	178795259	A/G			
rs189266	100020	178795270	C/G			
rs189267	100065	178795315	A/C			
rs170937	100086	178795336	C/G			
rs463263	101270	178796520	C/G			
rs463262	101276	178796526	G/T			
rs460454	101371	178796621	C/T			
rs460455	101376	178796626	C/G			
rs460505	101439	178796689	C/T			
rs931316	101820	178797070	C/T			
rs463431	102392	178797642	C/G			
rs461542	102602	178797852	A/G			
rs463557	102604	178797854	A/C			
rs191453	102896	178798146	C/T	0.15	0.19	0.139
rs2271212	189104	178884354	C/T	0.64	0.58	0.072
rs462009	189134	178884384	C/T			
rs2271211	189205	178884455	A/G			
rs396474	Not mapped	Not mapped	A/C			
rs428901	Not mapped	Not mapped	A/T	0.66	untyped	NA
rs452300	Not mapped	Not mapped	G/T			
rs670256	Not mapped	Not mapped	G/T			

TABLE 15

dbSNP rs#	Position in SEQ ID NO: 1	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs2278221	210	178695460	C/T	0.64	0.64	0.837
rs1650358	3608	178698858	C/G			
rs1643818	3609	178698859	C/G			
rs3733916	4318	178699568	C/T			
rs1624933	5593	178700843	A/G	0.73	0.75	0.447
rs1624857	5629	178700879	C/T	0.78	0.81	0.289

dbSNP rs#	Position in SEQ ID NO: 1	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs1624832	5639	178700889	A/G	0.44	0.47	0.423
rs1624829	5640	178700890	C/T	0.90	0.93	0.294
rs2161171	8943	178704193	A/C			
rs1530499	17968	178713218	A/G	0.39	0.36	0.499
rs888764	19887	178715137	A/G			
rs873987	21034	178716284	A/G			
rs4078699	21085	178716335	C/T	0.57	0.54	0.316
rs870311	21596	178716846	A/G	0.52	0.50	0.579
rs1643817	23379	178718629	A/C			
rs1643816	23432	178718682	A/C			
rs1650355	24007	178719257	A/C			
rs888763	26121	178721371	A/G	0.40	0.44	0.264
rs1862212	26273	178721523	A/T	0.56	0.53	0.529
rs1110514	26755	178722005	A/T	0.30	0.27	0.381
rs3797600	27411	178722661	C/T	0.55	0.54	0.840
rs3797602	27710	178722960	G/T	0.68	0.65	0.534
rs3797603	27842	178723092	C/T			
rs3776819	28379	178723629	C/T	0.45	0.47	0.662
rs252076	29603	178724853	C/T	0.46	0.46	0.986
rs252075	31232	178726482	C/G	0.36	0.34	0.666
rs252074	31504	178726754	A/G	0.35	0.33	0.604
rs252068	32583	178727833	C/G	0.47	0.48	0.648
rs252069	32794	178728044	A/G	0.27	0.26	0.640
rs194040	32840	178728090	C/T	0.31	0.30	0.734
rs252070	33044	178728294	C/T	0.61	0.55	0.157
rs3797606	33150	178728400	A/C	0.91	0.83	0.005
rs171667	33218	178728468	A/G	0.51	0.52	0.674
rs187539	33513	178728763	C/T	0.32	0.33	0.836
rs3836834	33959	178729209	- TATCA AACTAC CATGAA A			
rs252071	34486	178729736	A/G	0.30	0.30	0.942
rs252072	36289	178731539	C/T	0.50	0.49	0.684
rs252073	36570	178731820	C/T			
rs379589	38247	178733497	A/T	0.60	0.61	0.981
rs2052472	38477	178733727	A/C	0.06	0.06	0.856
rs2052471	38518	178733768	C/T	0.91	0.86	0.079
rs2052470	38529	178733779	C/T	0.82	0.83	0.828
rs2052469	38667	178733917	A/G	0.82	0.82	0.983
rs3797608	39781	178735031	C/T	0.06	0.06	0.969
rs3797609	39856	178735106	C/T	0.05	0.05	0.879
rs3822601	39927	178735177	C/T	0.07	0.08	0.838
rs153131	40506	178735756	A/G	0.76	0.76	0.981
rs751546	41869	178737119	C/G	0.91	0.92	0.526
rs2279979	42452	178737702	C/T	0.92	0.92	0.906
rs252060	44788	178740038	C/T	0.81	0.85	0.157
rs3797610	46059	178741309	A/C	0.18	0.16	0.593
rs194039	46846	178742096	A/G	0.39	0.49	0.005
rs168773	47712	178742962	A/T	0.37	0.43	0.098
rs252061	48796	178744046	C/T	0.19	0.15	0.164
rs187537	49441	178744691	C/G			
rs252062	49602	178744852	A/T	0.93	0.95	0.290
rs2431255	49723	178744973	A/C	0.23	0.19	0.201
rs3797612	50050	178745300	C/T	0.32	0.38	0.102
rs3797613	50171	178745421	C/T	0.23	NA	

dbSNP rs#	Position in SEQ ID NO: 1	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs614114	50477	178745727	C/T	0.48	0.51	0.423
rs252063	50818	178746068	C/T	0.60	0.51	0.011
rs252064	50833	178746083	C/T	0.51	0.56	0.265
rs252065	50881	178746131	A/G	0.22	0.18	0.175
rs450502	50882	178746132	A/G			
rs439252	51386	178746636	C/T			
rs252066	51534	178746784	C/T	0.18	0.16	0.451
rs457957	52317	178747567	A/G	0.67	0.68	0.728
rs3797614	52368	178747618	C/T			
rs423552	52970	178748220	A/G	0.89	0.91	0.398
rs398829	53023	178748273	A/G			
rs416646	53356	178748606	A/G	0.54	0.55	0.643
rs187450	53882	178749132	G/T			
rs337807	54553	178749803	C/T	0.49	0.59	0.009
rs337806	55475	178750725	A/C	0.11	0.10	0.889
rs1396438	55530	178750780	A/G	0.61	0.50	0.007
rs1396437	55691	178750941	C/T			
rs2411811	55848	178751098	A/C			
rs2898813	55879	178751129	C/G			
rs189256	56316	178751566	A/G	0.17	0.17	0.923
rs173072	56911	178752161	A/C			
rs337805	57320	178752570	A/G	0.27	0.25	0.582
rs191415	57391	178752641	C/T			
rs180045	57437	178752687	C/T	0.56	0.48	0.115
rs189255	57478	178752728	C/G	0.16	0.12	0.168
rs652766	57500	178752750	C/T	0.55	0.61	0.231
rs466750	59111	178754361	G/T	0.31	0.28	0.473
rs442406	59333	178754583	A/G	0.58	0.63	0.209
rs662407	59715	178754965	A/G	0.30	0.28	0.449
rs592971	59804	178755054	A/G			
rs457187	59851	178755101	A/G	0.23	0.21	0.402
rs459490	59929	178755179	C/T	0.20	0.19	0.708
rs459668	60052	178755302	C/T	0.21	0.20	0.821
rs462646	60240	178755490	C/T	0.44	0.41	0.460
rs458272	60359	178755609	G/T	0.22	0.20	0.524
rs463455	60381	178755631	A/G	0.23	0.22	0.629
rs675880	60456	178755706	C/T	0.65	0.67	0.564
rs810617	60724	178755974	C/G			
rs464156	60875	178756125	C/T	0.37	0.34	0.439
rs458083	60968	178756218	A/G			
rs467333	60978	178756228	C/G	0.11	0.11	0.902
rs465381	60998	178756248	C/T			
rs466363	61557	178756807	C/T	0.32	0.34	0.547
rs2457099	62091	178757341	C/T	0.43	0.43	0.974
rs463901	62645	178757895	C/T	0.39	0.43	0.342
rs465621	62943	178758193	A/C	0.59	0.64	0.195
rs463724	63131	178758381	A/T	0.09	0.07	0.539
rs465242	63145	178758395	G/T			
rs467419	63406	178758656	A/G	0.66	0.67	0.752
rs456135	63427	178758677	C/G	0.79	0.85	0.029
rs464536	63554	178758804	C/T	0.36	0.32	0.332
rs461898	63661	178758911	A/G	0.28	0.31	0.423
rs389558	64093	178759343	A/G	0.20	0.23	0.311
rs466752	64153	178759403	C/T	0.36	0.35	0.781
rs455655	64409	178759659	C/G	NA	0.72	NA
rs463435	64544	178759794	C/T	0.72	0.68	0.230
rs2174971	65257	178760507	C/T	0.56	0.51	0.142

dbSNP rs#	Position in SEQ ID NO: 1	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p- Value
rs1979979	65626	178760876	A/G	0.05	0.05	0.993
rs411804	65739	178760989	A/G	0.80	0.77	0.343
rs1623885	66392	178761642	C/T	0.84	0.84	0.819
rs1643811	66720	178761970	C/T	0.22	0.23	0.847
rs434430	69177	178764427	A/T			
rs187538	69336	178764586	G/T			
rs252067	69636	178764886	A/G	0.21	0.24	0.369
rs459319	69823	178765073	A/G	0.18	0.15	0.353
rs467289	69928	178765178	C/T	0.27	0.22	0.179
rs462644	70547	178765797	C/T	0.60	0.61	0.609
rs458752	70633	178765883	C/T	0.18	0.15	0.271
rs708320	71805	178767055	A/C			
rs457954	72181	178767431	C/G	0.72	0.72	0.882
rs2411810	72200	178767450	C/T	0.29	0.30	0.630
rs3084687	72474	178767724	-AT	0.13	0.11	0.509
rs69638	72567	178767817	C/G	0.54	0.57	0.440
rs455452	72973	178768223	A/G	0.60	0.58	0.499
rs464850	73468	178768718	A/G	0.10	0.09	0.839
rs431472	73889	178769139	A/G	0.35	0.27	0.025
rs2411809	75730	178770980	C/T			
rs2457094	75970	178771220	A/G	0.71	0.70	0.792
rs2457095	76114	178771364	A/G	0.75	0.76	0.602
rs2261740	76342	178771592	C/T	0.36	0.36	0.924
rs1109180	76449	178771699	A/G			
rs1109179	76465	178771715	C/T			
rs1109178	76791	178772041	A/C	0.45	0.42	0.420
rs456909	78042	178773292	A/G	0.53	0.51	0.598
rs469124	80758	178776008	A/G			
rs468039	80778	178776028	C/T			
rs467017	81356	178776606	A/C	0.34	0.35	0.762
rs469290	81576	178776826	A/G	0.49	0.44	0.223
rs469090	81689	178776939	C/T	0.83	0.84	0.883
rs469568	81759	178777009	G/T	0.36	0.30	0.115
rs468386	81950	178777200	C/G			
rs469349	82562	178777812	A/C			
rs469099	83591	178778841	C/T	0.65	0.67	0.560
rs456868	83700	178778950	A/G			
rs465389	83821	178779071	C/G			
rs463892	83842	178779092	C/G			
rs468548	83923	178779173	G/T			
rs654612	83929	178779179	A/C			
rs468542	84021	178779271	C/G			
rs469262	84175	178779425	C/T	0.45	0.43	0.762
rs708323	84417	178779667	A/G	0.74	0.74	0.899
rs469089	84747	178779997	C/G			
rs469396	85746	178780996	C/G	0.39	0.42	0.569
rs468723	86129	178781379	C/T	0.36	0.34	0.573
rs467604	86335	178781585	A/G	0.35	0.36	0.763
rs338874	87315	178782565	C/G	0.42	0.40	0.564
rs338875	87648	178782898	A/G	0.46	0.45	0.701
rs1385803	87764	178783014	A/C			
rs1385804	87770	178783020	C/G			
rs338876	88221	178783471	C/T	0.41	0.44	0.580
rs189803	90474	178785724	A/C			
rs452215	91148	178786398	G/T			
rs641170	91150	178786400	G/T			
rs584398	91160	178786410	G/T			

dbSNP rs#	Position in SEQ ID NO: 1	Chromosome Position	A1/A2 Allele	F A2 Case AF	F A2 Control AF	F p-Value
rs385330	91733	178786983	C/T			
rs429538	91772	178787022	A/C			
rs371229	91785	178787035	C/T			
rs460874	93140	178788390	A/T	0.73	0.75	0.550
rs646121	93148	178788398	A/T	0.93	0.92	0.697
rs468262	96080	178791330	A/G			
rs467863	96157	178791407	C/G			
rs191434	96313	178791563	A/C			
rs2054782	96759	178792009	C/T	0.43	0.40	0.473
rs468499	97026	178792276	A/C			
rs180287	97320	178792570	C/G			
rs338877	97732	178792982	A/T	0.04	0.04	0.928
rs650665	98713	178793963	C/G			
rs193419	99707	178794957	A/C			
rs180288	99959	178795209	C/G			
rs186834	100009	178795259	A/G			
rs189266	100020	178795270	C/G			
rs189267	100065	178795315	A/C			
rs170937	100086	178795336	C/G			
rs463263	101270	178796520	C/G			
rs463262	101276	178796526	G/T			
rs460454	101371	178796621	C/T			
rs460455	101376	178796626	C/G			
rs460505	101439	178796689	C/T			
rs931316	101820	178797070	C/T			
rs463431	102392	178797642	C/G			
rs461542	102602	178797852	A/G			
rs463557	102604	178797854	A/C			
rs191453	102896	178798146	C/T	0.06	0.06	0.929
rs2271212	189104	178884354	C/T	0.66	0.56	0.012
rs462009	189134	178884384	C/T			
rs2271211	189205	178884455	A/G			
rs396474	Not mapped	Not mapped	A/C			
rs428901	Not mapped	Not mapped	A/T	0.61	0.72	0.002
rs452300	Not mapped	Not mapped	G/T			
rs670256	Not mapped	Not mapped	G/T			

[0239] Allelotyping results were considered particularly significant with a calculated p-value of less than or equal to 0.05 for allelotype results. These values are indicated in bold. The allelotyping p-values were plotted in Figure 1 for the discovery cohort. The position of each SNP on the chromosome is presented on the x-axis. The y-axis gives the negative logarithm (base 10) of the p-value comparing the estimated allele in the case group to that of the control group. The minor allele frequency of the control group for each SNP designated by an X or other symbol on the graphs in Figure 1 can be determined by consulting Table 13. For example, the left-most X on the left graph is at position 178695460. By proceeding down the Table from top to bottom and across the graphs from left to right the allele frequency associated with each symbol shown can be determined.

[0240] To aid the interpretation, multiple lines have been added to the graph. The broken horizontal lines are drawn at two common significance levels, 0.05 and 0.01. The vertical broken lines are drawn

every 20kb to assist in the interpretation of distances between SNPs. Two other lines are drawn to expose linear trends in the association of SNPs to the disease. The light gray line (or generally bottom-most curve) is a nonlinear smoother through the data points on the graph using a local polynomial regression method (W.S. Cleveland, E. Grosse and W.M. Shyu (1992) Local regression models. Chapter 8 of Statistical Models in S eds J.M. Chambers and T.J. Hastie, Wadsworth & Brooks/Cole.). The black line provides a local test for excess statistical significance to identify regions of association. This was created by use of a 10kb sliding window with 1kb step sizes. Within each window, a chi-square goodness of fit test was applied to compare the proportion of SNPs that were significant at a test wise level of 0.01, to the proportion that would be expected by chance alone (0.05 for the methods used here). Resulting p-values that were less than 10^{-8} were truncated at that value.

[0241] Finally, the exons and introns of the genes in the covered region are plotted below each graph at the appropriate chromosomal positions. The gene boundary is indicated by the broken horizontal line. The exon positions are shown as thick, unbroken bars. An arrow is placed at the 3' end of each gene to show the direction of transcription.

Example 5

Effect of ADAMTS2 Polypeptides on Biosynthesis of Type II Collagen in Patients with OA

[0242] To investigate the effect of *ADAMTS2* polypeptides on Type II collagen biosynthesis and processing, human articular cartilage from OA patients undergoing joint replacement is harvested, dissected and maintained as described by Nelson *et al.* (1998) *supra*. Type II procollagen levels in osteoarthritic patients and autopsy controls is determined by radioimmunoassay (RIA) as previously described. Allelic variations (*e.g.*, rs398829) are determined for the OA patients and controls by genotyping (See Examples 1 and 2). As type II procollagen is processed by *ADAMTS2*, increased levels of Type II procollagen in individuals with the allelic variation associated with OA demonstrates that this variation leads to reduced procollagen processing activity and ultimately to OA.

Example 6

Effect of ADAMTS2 Polypeptides on Type II Collagen Processing Activity

[0243] To investigate the effect of ADAMTS2 polypeptide variants on ADAMTS2 collagen processing activity, recombinant polypeptides encompassing the ADAMTS2 variation of SEQ ID NO: 2 at position 733 and a wild-type ADAMTS2 polypeptide are expressed in cell lines such as chondrocytes. Since the allelic variation of ADAMTS2 at position 733 of SEQ ID: NO: 2 will prevent the conversion of the ADAMTS2 pro-enzyme to the catalytically active enzyme, processing of ADAMTS2 pro-enzyme is monitored by SDS-PAGE analysis followed by Western Blotting using antibodies to ADAMTS2 and

methods common to someone skilled in the art. Reduced levels of pro-enzyme cleavage are apparent by the increased levels of immunopositive protein of higher molecular weight than of the cleaved active protein.

Example 7

In Vitro Production of Target Polypeptides

[0244] cDNA is cloned into a pIVEX 2.3-MCS vector (Roche Biochem) using a directional cloning method. A cDNA insert is prepared using PCR with forward and reverse primers having 5' restriction site tags (in frame) and 5-6 additional nucleotides in addition to 3' gene-specific portions, the latter of which is typically about twenty to about twenty-five base pairs in length. A Sal I restriction site is introduced by the forward primer and a Sma I restriction site is introduced by the reverse primer. The ends of PCR products are cut with the corresponding restriction enzymes (*i.e.*, Sal I and Sma I) and the products are gel-purified. The pIVEX 2.3-MCS vector is linearized using the same restriction enzymes, and the fragment with the correct sized fragment is isolated by gel-purification. Purified PCR product is ligated into the linearized pIVEX 2.3-MCS vector and *E. coli* cells transformed for plasmid amplification. The newly constructed expression vector is verified by restriction mapping and used for protein production.

[0245] *E. coli* lysate is reconstituted with 0.25 ml of Reconstitution Buffer, the Reaction Mix is reconstituted with 0.8 ml of Reconstitution Buffer; the Feeding Mix is reconstituted with 10.5 ml of Reconstitution Buffer; and the Energy Mix is reconstituted with 0.6 ml of Reconstitution Buffer. 0.5 ml of the Energy Mix was added to the Feeding Mix to obtain the Feeding Solution. 0.75 ml of Reaction Mix, 50 μ l of Energy Mix, and 10 μ g of the template DNA is added to the *E. coli* lysate.

[0246] Using the reaction device (Roche Biochem), 1 ml of the Reaction Solution is loaded into the reaction compartment. The reaction device is turned upside-down and 10 ml of the Feeding Solution is loaded into the feeding compartment. All lids are closed and the reaction device is loaded into the RTS500 instrument. The instrument is run at 30°C for 24 hours with a stir bar speed of 150 rpm. The pIVEX 2.3 MCS vector includes a nucleotide sequence that encodes six consecutive histidine amino acids on the C-terminal end of the target polypeptide for the purpose of protein purification. Target polypeptide is purified by contacting the contents of reaction device with resin modified with Ni²⁺ ions. Target polypeptide is eluted from the resin with a solution containing free Ni²⁺ ions.

Example 8
Cellular Production of Target Polypeptides

[0247] Nucleic acids are cloned into DNA plasmids having phage recombination sites and target polypeptides are expressed therefrom in a variety of host cells. Alpha phage genomic DNA contains short sequences known as attP sites, and *E. coli* genomic DNA contains unique, short sequences known as attB sites. These regions share homology, allowing for integration of phage DNA into *E. coli* via directional, site-specific recombination using the phage protein Int and the *E. coli* protein IHF. Integration produces two new att sites, L and R, which flank the inserted prophage DNA. Phage excision from *E. coli* genomic DNA can also be accomplished using these two proteins with the addition of a second phage protein, Xis. DNA vectors have been produced where the integration/excision process is modified to allow for the directional integration or excision of a target DNA fragment into a backbone vector in a rapid *in vitro* reaction (Gateway™ Technology (Invitrogen, Inc.)).

[0248] A first step is to transfer the nucleic acid insert into a shuttle vector that contains attL sites surrounding the negative selection gene, ccdB (*e.g.* pENTER vector, Invitrogen, Inc.). This transfer process is accomplished by digesting the nucleic acid from a DNA vector used for sequencing, and to ligate it into the multicloning site of the shuttle vector, which will place it between the two attL sites while removing the negative selection gene ccdB. A second method is to amplify the nucleic acid by the polymerase chain reaction (PCR) with primers containing attB sites. The amplified fragment then is integrated into the shuttle vector using Int and IHF. A third method is to utilize a topoisomerase-mediated process, in which the nucleic acid is amplified via PCR using gene-specific primers with the 5' upstream primer containing an additional CACC sequence (*e.g.*, TOPO® expression kit (Invitrogen, Inc.)). In conjunction with Topoisomerase I, the PCR amplified fragment can be cloned into the shuttle vector via the attL sites in the correct orientation.

[0249] Once the nucleic acid is transferred into the shuttle vector, it can be cloned into an expression vector having attR sites. Several vectors containing attR sites for expression of target polypeptide as a native polypeptide, N-fusion polypeptide, and C-fusion polypeptides are commercially available (*e.g.*, pDEST (Invitrogen, Inc.)), and any vector can be converted into an expression vector for receiving a nucleic acid from the shuttle vector by introducing an insert having an attR site flanked by an antibiotic resistant gene for selection using the standard methods described above. Transfer of the nucleic acid from the shuttle vector is accomplished by directional recombination using Int, IHF, and Xis (LR clonase). Then the desired sequence can be transferred to an expression vector by carrying out a one hour incubation at room temperature with Int, IHF, and Xis, a ten minute incubation at 37°C with proteinase K, transforming bacteria and allowing expression for one hour, and then plating on selective media. Generally, 90% cloning efficiency is achieved by this method. Examples of expression vectors

are pDEST 14 bacterial expression vector with att7 promoter, pDEST 15 bacterial expression vector with a T7 promoter and a N-terminal GST tag, pDEST 17 bacterial vector with a T7 promoter and a N-terminal polyhistidine affinity tag, and pDEST 12.2 mammalian expression vector with a CMV promoter and neo resistance gene. These expression vectors or others like them are transformed or transfected into cells for expression of the target polypeptide or polypeptide variants. These expression vectors are often transfected, for example, into murine-transformed a adipocyte cell line 3T3-L1, (ATCC), human embryonic kidney cell line 293, and rat cardiomyocyte cell line H9C2.

[0250] Modifications may be made to the foregoing without departing from the basic aspects of the invention. Although the invention has been described in substantial detail with reference to one or more specific embodiments, those of skill in the art will recognize that changes may be made to the embodiments specifically disclosed in this application, yet these modifications and improvements are within the scope and spirit of the invention, as set forth in the claims which follow. All publications or patent documents cited in this specification are incorporated herein by reference as if each such publication or document was specifically and individually indicated to be incorporated herein by reference.

[0251] Citation of the above publications or documents is not intended as an admission that any of the foregoing is pertinent prior art, nor does it constitute any admission as to the contents or date of these publications or documents. U.S. patents and other publications referenced herein are hereby incorporated by reference.

Nucleotide and Amino Acid Sequence Embodiments

[0252] Table A includes information pertaining to the incident polymorphic variant associated with osteoarthritis identified herein. Public information pertaining to the polymorphism and the genomic sequence that includes the polymorphism are indicated. The genomic sequences identified in Table A may be accessed at the http address www.ncbi.nih.gov/entrez/query.fcgi, for example, by using the publicly available SNP reference number (e.g., rs398829). The chromosome position refers to the position of the SNP within NCBI's Genome Build 34, which may be accessed at the following http address: www.ncbi.nlm.nih.gov/mapview/map_search.cgi?chr=hum_chr.inf&query=. The "Contig Position" provided in Table A corresponds to a nucleotide position set forth in the contig sequence (see "Contig Accession No."), and designates the polymorphic site corresponding to the SNP reference number. The sequence containing the polymorphisms also may be referenced by the "Nucleotide Accession No." set forth in Table A. The "Sequence Identification" corresponds to cDNA sequence that encodes associated target polypeptides (e.g., ADAMTS2). The position of the SNP within the cDNA sequence is provided in the "Sequence Position" column of Table A. If the SNP falls within an exon, the corresponding amino acid position (and amino acid change, if applicable) is provided as well. Also, the

allelic variation at the polymorphic site and the allelic variant identified as associated with osteoarthritis is specified in Table A. All nucleotide and polypeptide sequences referenced and accessed by the parameters set forth in Table A are incorporated herein by reference.

Table A

RS_ID	Chromosome	Chrom Position	Contig Accession No. [1]	Contig Position	Nucleotide Accession No. [2]	Sequence Position	Amino Acid Position	Locus	Locus ID	A [3]	Allelic Variability	OA Assoc. Allele
rs398829	5	178748273	Hs5_77500_34:3	1729878	NM_014244	coding-nonsynon	V245I	ADAMTS2	9509	R	[G/A]	G

[1] Contig Accession Number which can be found in the NCBI Database:
http address: www.ncbi.nih.gov/entrez/query.fcgi

[2] Sequence Identification or Nucleotide Accession Number which can be found in the NCBI Database:
http address: www.ncbi.nih.gov/entrez/query.fcgi

[3] "A" column is the sequence orientation ("F" is forward, "R" is reverse).

[0253] The following is a genomic nucleotide sequence for an *ADAMTS2* region. The following nucleotide representations are used throughout: "A" or "a" is adenosine, adenine, or adenylic acid; "C" or "c" is cytidine, cytosine, or cytidylic acid; "G" or "g" is guanosine, guanine, or guanylic acid; "T" or "t" is thymidine, thymine, or thymidylic acid; and "I" or "i" is inosine, hypoxanthine, or inosinic acid. Exons are indicated in italicized lower case type, introns are depicted in normal text lower case type, and polymorphic sites are depicted in bold upper case type. SNPs are designated by the following convention: "R" represents A or G, "M" represents A or C; "W" represents A or T; "Y" represents C or T; "S" represents C or G; "K" represents G or T; "V" represents A, C or G; "H" represents A, C, or T; "D" represents A, G, or T; "B" represents C, G, or T; and "N" represents A, G, C, or T.

ADAMTS2 Genomic Sequence (SEQ ID NO. 1)

```
>5:178695251-178884700
1      gggcctcgaa ccagcaactgc ccagcgctgg gaagacagga gaccacaccc caagtggcctt
61     tgacacaggc cgtgtccctc ttaaggcaca gaggagaagt gggcagccag ggctggggat
121    cctagggtgg cccctctgtg ccccacccctt ccccaaggcca cttacacgtg gccagtctca
181    tgggccacca caaacgcgtga ggagaagccR tcctcatgtt tcagggtgca gctgcggacc
241    ggatggcaca tgccggtgac aggagcatag cctggggagga gacaagaggc ggctccagat
301    gctgccatag cctggccggg aaggtagggc ctggcttaac tccccggcgc tgcttcctct
361    caaggccccca atgccccttc taccacaggc aactgcccccc ggctggcaca gctcagaaga
421    caccacagca gagagctcat agcctgtgac atctggctga cagcaggccc ccagcccgcc
481    accacacaag ccccggtggcc gtctgcctg taacagccac actgcagctg gggccctctg
541    cttagtcaga tgtcacctgc tatggcttgt ctgacacctg actttcctgc ctggcggttct
601    cttagtcaga acctgcctct gccttcactg agtggccctt gaggcagcacc
661    tgcgtccaca cctctctctg ctagaggcat cctgccccctc gtctgcctgg tatcacaccc
721    tccaaaggcca cctctctctgg gggcggtttc ttgggtctcc ccgtgtttct ctggggctct
781    tgagacccta gctagacactg tttccataat gccgcagaag gctgcaaagc tgttcagccg
```

841 atgggccaca ctaccctgca gaatatccca gcagggggct ggctggtgcg gactgggagg
901 tcagagccgc cgtagaagat gagggcagtgc gcctacgaca gcccaggccc aaaaaaaaaaaaa
961 tggcagaagt tcatgacgca gagacccccc gtcaaactga ggaagatggta taatgagacc
1021 tattcaacttc cccagctccgt gtctcagcag agggaggccc cttggaaaga tagggaaagt
1081 cctacaggc tatgggtgcc acacgaggc tctgcttgc cccctggag tcttcttaga
1141 ggcttgccaa gcatggagcc ctgactttca gaaccccaac ctgcctcac tcctgcccatt
1201 ccccaacttag agacccttag gccaacgcgtt ccacagagaa gggaggagg gccactggg
1261 ccccggtga gctggcggtg tacgcccataaaaatagaag catcgatc taacccttgt
1321 ctttccctgt tgtaataaaac agggaaatct gcatttttg ctctgcagat tttgagacga
1381 gctggagtc tggcgcctg agtggAACgt tgccatctcc tcttgcctt ttgtatgcca
1441 ggcttgccaa tgcatctgcc tccgtggag gctttccatt catccactc agcattacag
1501 gagcctgca gatctgaaag cgcatctta tggtttctt tggtttctt cctcttattt
1561 tactttgtgg ctgtttctc tctccgaaat cttatttagat agatgtgcc tcctcgatgg
1621 agggctgtc tcctaacttt tctctcatgt taatctttt gtctttgtc cccattggga
1681 gatggcccttgc acttttattt tcagcttgc atttttttt tttttttt agatggagtc
1741 tcaactgtcc gtccaggctg gagtgcatgt gtgtgatctc ggctactgc aacccccc
1801 tcctgggttc agtccatctt actgcctcgtt cctctcgat agctggact acaggcaccc
1861 gcaaccatgc tcagctaatt tttgtatTT tactagagac gggatttcac catgttagcc
1921 agggatggctc ccatctcctg acctcggtac ctgcccgcct cggcctccca cagtgctgg
1981 attacaggcg tgagccaccg cgccccgtct acctttcgaa ttttttactt ccccatctt
2041 atccataattt ttatacaaaat ggtttggatg tctctaattt ttcatagaaa actcttaggg
2101 tttatggatg tattttotca aacattctgg gaataactact tttaaaaaca tgctttctt
2161 gtctctaaat tatctctattt ttctctttaga gcccattttc aaactttttt ttttggat
2221 catcttggtc catcttttc atgctgtaa gatttcttgc ttttttttt tttttttt
2281 tctggtaacc ctggcgttc ctttcatttt ttcttttttt tttttttt tttgagacag
2341 agtcttgc tggcccgag gctggagtgc agtggcacaat tctcgctca ctgcaacc
2401 tgccctccgg gttcttagta ttttctgc tcaagccttgc gatgtggc gattacaggc
2461 acctgcact acactggta atttttgtgt tttttagtaga gatgtggc aggctggta
2521 cgaactcctg acctcaagtg atccggccgc ctggcctcc caaatgtctg ggattacagg
2581 agtgagccac cacacccacg caatttcaca tttttaaatg agggatcagg tcaacttaata
2641 taaaagctgt gtcggttcc cctcttaggg gatagctctg ctcccaattt gtgcctccc
2701 caagacaaaaa cagccaaacag cactctcattt acatggtcag ggcaacacct ggctgegctg
2761 tggttgaggg gagctatcca ctccctaccc agacatggta gttttagtct ttttacacta
2821 gctatgttat taggtgcgtc ttgatattt attgtgttt taaaatttcc tttttttt
2881 ttcttgttaag acaaggccctc gctctgtcac ccaggctgg gtcagtggt gtgatcatag
2941 ctcacttctt gggctcaaac aatcttccca cctcagggttcc cccaggctgg gggactacag
3001 gcatccacca ccatggcccg cctcattttt ttaattgtat ttttccaaatg tctagtgg
3061 ttgagtgccct tttcatatgt ttatgtctca ttcatatttt gtcaagagcc tgataagagt
3121 ctccctggccca ttatctgact ggattgtttt tattatctt attatgattt gtaggcgtc
3181 ttacacatact tcagatataa atctttttt ataaatattt ccccaactgg attagcgtc
3241 cttaatgca caaaactttt aatgtgtca ctatattatca gtttttccca ttgagttcat
3301 ggtttggttt gtttttttgc tctttttttt tttttagagaca gatgtctcgca ctgtctccc
3361 ggctggagca aagtggtgcg atctcgactc attgcaaaactt ccgcctcccg ggttcacacc
3421 attctccgtc ctccggccctc cggtagctg ggactacagg cggccggccac catggccctgc
3481 taatttttt gtatTTTTT tagagacggg gtttccattt gtagccaaatg atggctcg
3541 tctccgtacc tcgtgtatctt cctccctccg actcccaaaag tgctggattt acagggtgt
3601 accaccgSSc cccggccctt gtcctttttt aaagaaaaaca tctcttactc caagaacatg
3661 gagccatcca cctacgttat tttttccat tttttgggtt tttttttt gacctaaaaat
3721 ccatgtgaga tgctctgtt ctatagagtg agtggagggtt acagccctcc tttttttt
3781 aattggccccc acaccatTT ttaggctttt ttcctccctt tctcttggaaat gtcactttt
3841 tgacacatta gctcccttat acacataggt ccaattcagg acactggatt ctcttctaat
3901 gatgtatTT tccatcgatg cactggccca caccctcacc cagctaaqgt ggccacqgg
3961 ttccctggcc tgcacttcc cccggggctg ggcaaggct cccggggccc cttgcattggc
4021 caggggctca ttccatgtt cacccctctg tggccctgc tgggtggc tcaccatttgc
4081 tgccggaaagg cccaaagtcc tgccgtgtga ggaagatggc gtgatcggtt tattcatgt
4141 ggccctgtc tggcttctgc tggaggttggccc cccagccggca gacattctcc aggctctg
4201 aggggttccc gatctcgatg agggctcatgg actgcagggg gatggagaga aatggaaag
4261 agagggttggc gcctgggtggc ccagggtgggg gccggccagg ctgtgtgtt gacagacYcc
4321 atccacttggg aatctgtttcc aggtggtaca aaggccctgc cccggccactc tcaggaccc
4381 ctcccaactgg gctccagttt ttgacaggga agaaaaaggat tcccgactcactcacc
4441 gtgagttcc acacaaagtcc acccgaggat tcaagggttgg gtcagaagcc aagctctac
4501 tcatctaaagg cctcgcatag tccaccgtc tacccttccat ccagaagttcc ctgtgtttt
4561 tctatggggc acctgttccg tgcaaggccct ccacagaaggaa caaatgtcat gtggaccc
4621 ccccaaagag ctcacagttt tggagggc cgtggcagggg gaggcggaaa ttccagtgt

4681 gtgtgcaagg ccagaggaag cagatcacct tatacggtcc ccccggtgtga aggtacttgc
4741 agaaaaccac acaatggagg aaaagaaaaca gcaactgaca tattttgtgg gtgc当地
4801 aatcgcagta ttgcccataa aagtaatggc aaagacttt gcatcaaccc aatagaaaa
4861 cagccatatac aagaacaagg gcatacacag atgattccca caggaactat gcatggaaag
4921 cccccggaa ccatccatt ccattttgtt tcatgacgca cacctgacag cttaggtggc
4981 ctcgggaccg ctcgtctgc acactcttc tgccatcatg catcagcagg cactccatt
5041 tagacttcag gactaagaag cctgtcttg acaggaaact gagactcaga gaagtaaga
5101 agcttgcaca ggtcatggaa ctaggaagta ggagaattga gatcgatcc caagaaggct
5161 gaaacccaaa tcttagattca aatattcattt caaaagcaag gtttctgtga gccc当地
5221 ctcacagcag gatctttgtt ttctgttcc accatagact ccaaacaccg aagtatcc
5281 ctccaagaga accgagagcc cctctgagtc ctggcatgt tctttcccc ttaaggatt
5341 tatctggcc taaaatagag atgtgcattt cttgaaagct acgtatgcag gtgttttcc
5401 tcttgcataa atgtatccac aaagtagagc cacgtattt tctttgtgtc tataaaaac
5461 aaaaagtgg ccgtgcgcgg tggctcacac ctgtaatccc agcactttg gaggctgagg
5521 caggtggatc acYtgaggtc aggagttga gaccagctg gccaacatRg taaaacccYR
5581 tctctactaa aaatgcaaaa cttagtctgg cttgtggcg taatcccagc aacctggag
5641 gctgaggaag gagaatcgct ccactgcaact ccagcctagg caacaaaggg ttttctgtcc
5701 aagccaggtt aagtttgcgg aagccaggtt gttccagaag aaggagcac aattcaaggc
5761 taggttaagac aagtcttc tcccaggcc agtcaatgac tgacaagat gcagaccaat
5821 tcccaagtttca caagcacaga gtaatttgcgg ttttctgtcc tgggatataa
5881 actagcttc tctaacaagg ttttctgtcc ttttctgtcc tgggatataa
5941 gataactcat ttgttcctg tgactaagac ttttctgtcc ttttctgtcc tgggatataa
6001 ttttctgtcc ttttctgtcc tgggatataa
6061 ctggaggagg gcaaggggg actagcttc gccaatgagc ttttctgtcc tgggatataa
6121 tcttaatttc cattgagacc ttttctgtcc tggcacttcc ttttctgtcc tgggatataa
6181 taaagacagt ggctgcttg tcacccatgg ttttctgtcc tggcacttcc tgggatataa
6241 ccaaccacca ttgaacatac ttttctgtcc tggcacttcc tgggatataa
6301 tgcaaagtgt gtttctgtcc ttttctgtcc tgggatataa
6361 ggagtggta tctgcatcac ttttctgtcc tggcacttcc tgggatataa
6421 aggttagtac cgccctctggg aagaaatgtc ctttctgtcc tgggatataa
6481 aaaggaaact cgggggtggat ttttctgtcc tgggatataa
6541 ctcgctgaac ttaagattgg ttttctgtcc tgggatataa
6601 ttttctgtcc tgggatataa
6661 ttttctgtcc tgggatataa
6721 ttttctgtcc tgggatataa
6781 ttttctgtcc tgggatataa
6841 aaacagttga ttttctgtcc tgggatataa
6901 ggattgcca ggtggta ttttctgtcc tgggatataa
6961 ttttctgtcc tgggatataa
7021 gctagaagag tggactttga ttttctgtcc tgggatataa
7081 gatatgtta ttaccctgt ttttctgtcc tgggatataa
7141 ccctacccca gaaatatgtt ttttctgtcc tgggatataa
7201 atcaaaaacc aaacaagaat ttttctgtcc tgggatataa
7261 tccaaatccag aggtcagca ttttctgtcc tgggatataa
7321 gggccctgtc gcaactgtgag ttttctgtcc tgggatataa
7381 ctgggctca gggaaagggg ttttctgtcc tgggatataa
7441 agggagttct gtccaaagat ttttctgtcc tgggatataa
7501 acggccctt ccgtatggcc ttttctgtcc tgggatataa
7561 cgtaaactt cccctcttt ttttctgtcc tgggatataa
7621 ggccaatccc agcaccacgc ttttctgtcc tgggatataa
7681 aaagcgggccc aaggagtagg ttttctgtcc tgggatataa
7741 ctgcactaca ggagaccccc ttttctgtcc tgggatataa
7801 gactggggct gcaactgtcat ttttctgtcc tgggatataa
7861 tcaggtggtc acggcgggcc ttttctgtcc tgggatataa
7921 tggggccgtc acccatgtca ttttctgtcc tgggatataa
7981 agacaacaag agaaggcacc ttttctgtcc tgggatataa
8041 gtcccgccca ccgcaggggca ttttctgtcc tgggatataa
8101 gtcgcctgca actgacgggt ttttctgtcc tgggatataa
8161 aattctgagt ttttctgtcc ttttctgtcc tgggatataa
8221 ttttctgtcc tgggatataa
8281 atctggggc agctgatgt ttttctgtcc tgggatataa
8341 cctggccaaag ttttctgtcc tgggatataa
8401 ggaccagagg caggaatgag ttttctgtcc tgggatataa
8461 ggcaggctgg ggttgtgcag ttttctgtcc tgggatataa

8521 cgggaggatg ctagaaaccc aagcgcctac aggagcaagg caggtgacaa gtcaatgagg
8581 caggctggtt gcacgaaaag taagtccatc ttgtgtttc ccagtttga aggagaaaagg
8641 gtacagagga atatttcctt attagaatag agtgggggtc ggtgataagg agtggcgagg
8701 actgtggcaa acccaggtgc ccaggcacag ctgaaggagg caacatccgc ccagtgtgg
8761 caggaacaga ggcccagagc tgctggatca gatttctcag gtcaatcgtg aaatctggat
8821 tctacagaca cctacctaatt agtaaacact ggctcaaatt gaaatcgtt atgcagacaa
8881 aacaacgcatt cctccccgtc accactttt tccagaaccc tggataaga gccctgcagt
8941 ggktagcccc ttcttccaatt cctcgacatc ggatccactc tggccgatca cctcttcct
9001 cccagatgct ggcttctgtg gttcagcctg gtttccacgc ccgtcttagc agccaaagt
9061 atcttgctgc tgcggcaaga caatttttg aaatattgtt ttgtgtttt gttcttgctt
9121 cgctttgttc tcggcccaat ccaagaaggc agcctgccaac ttcgttctc caaccccaag
9181 gaggttgtgc tgcagggtcc agagctaaag gggtctatga ccccttctg gtttgggg
9241 tctgaggaga ctccctgagg agggggcaga ggggagggtc ggggtgcatt ggcacaggga
9301 gggaaagctg cactggggcc aggggagccc tggggaggcag gacttcagga gacatgagag
9361 tgtgttagac cctggggagcc tgggtcccaat caggggggtac cccaagaagt agaggagagg
9421 gttaccttag tgcggggca cccaggact caacctgcca ctgggagctt aactgacaag
9481 ccacagact gcttgcctt cctcttc tctgccaacatc tccaaaggtcc
9541 cagcttatct ctgacggcga cagaggagcc cagggggcaga gatctgttcc ttgctgggtc
9601 tggagcatc aagggtctatg ggttccatc gaggcccaat cattgaccctt gtgcccagag
9661 gccaggccaa ggtctgtgtg cattttttt acagccctt gggtcagtga ggaaggagtg
9721 gggatagaag gggacagtgt ggactgaga tgcattgcagg tttgtgttgc gggacgttag
9781 gagggtcagg acctgggtgc aggagctgagg aggcatttcagg gctgcaggc gggagcaatg
9841 aggccagagac ggcaccatgt tctggctttt gggattggta gcaaggcaga agaggccgg
9901 cggaaagggg acggcaatgt ccagaggaag agcgtggggt ccgcagtgc gaaacatgg
9961 tcacggcaat gtcttcctt gccagcttgc tcgttggaa gagctggctt gtttgcatt
10021 gggagctctc ccacttctgg ggggtcagcc cttgtgggtt atcagtcatac agccgtatgt
10081 gggacccatg ttaggccccat ggagcacagg ttcgaaagga ctgaaaactg gctgaagtgg
10141 gtttgcctgg tggcagcaccc caaaaacaacc atccgcacac tcaggtcat gcaccccaaga
10201 gcccccttc ctccacccac tcctccact ccatgagagc tacaggcgc cttgggttct
10261 gcatgcctt aaggactctt cacaaaaactt aactgtatc caggcagaga catcggggca
10321 gctggacagg tctcgccccc aggagggaaat tctgggttgc ccgtggacag ggcaggccc
10381 tcgctactgg aggccatgttgg aaccccaaggg aagagccaaat ctcaggtctc tccgctgaga
10441 cctccagccca aggataactgt cctgacatgc agtgaacttcc cctgtgggtc acttgggtc
10501 cagcgggacc acctgttggg atgcacatc ctttgcctt ctgcacactg gcctgcattc
10561 aggaggccag atgggtccct gtcgttggg gcaccccaac cccacccatc ctgtgcctt
10621 gctatgtatc aagectggga actgtgttcc ccctgagcga gaagaagaca catttctcat
10681 atatggatct cctccagaag tgagggaggg tgcactgtt gtttccactg aggagtctt
10741 caccagctcg ccctggatcat ctacgttccct cccagtcctg aatgcacactg atttccact
10801 gctgtctcggt ctctgcacgg ctcttcatttggg aaagtctgtc tttgtgtactc caccgttcc
10861 tagctgtgtc gagggggggg ccaggactgtc gccgttagcac cctgggtctc gcaagggggg
10921 tggctgcagg accaggccagg gacttttacag ggcaggccat ccctcccaagg ggcaggcccg
10981 ggcacccaggc ctccagagca gggggcggtt gatgggtgtt gatgaccggg gatggacatt
11041 tcgaggccctt ggggtcgtcat ggttagctgtca ggttccacctt cctgatgttgc gcaaggccct
11101 tgagggcaga ctcttcacag cggcagcagc tcccaactt cccaggttca gcctaggccc
11161 ccagaagagc caggcgggag aaggctgcag tagggctgtt gatgaaatgttcc agggggccat
11221 ccagcccaagg atgcaggagg agtcttgcgtt gggctgaccc tagtcagag cacgggaaca
11281 gatgagcttca aagaggcttgc atccgttccctt ctacaggggc cccaggacag tcgttacccc
11341 tcacggggccat cctgtttggg ggtgacatcg accactgcag acatgggttc
11401 ccaactgggc tccactgccc gagtgcgttgc gggcccttgg ccaggctggc tttccagag
11461 tgcagggtcc tccttggccat tcatgatgtc ccacatgttgc ccacatggag acccttcttgc
11521 ctatgtgttgc tcaacagatgc gcccagatgc aggaggcaga cactgttca gggcccttgg
11581 gccccttccatc agcaggccagg ggggttggg cccaaaggagg atgggttttcc tccgggttgc
11641 agcctgttttgc gtcgttggggccat cccgggttggg aaagggttgc aggacatgtcc taaaagacc
11701 agggatccca ccatgagggttgc gagggggcag atctgagggttgc cttaggaggaaatgagctt
11761 atggccctca ctttgggttca gaagggttact ttgtgttcaat ttttttttttgggggg
11821 gtttggccat ggcacgggttgc agggaaacactt gcccggccctt gaaatgttgc cccaggcttgc
11881 aggggtctgg gtaacccac gggcccttggg gaaatgttgc gaggttgggg gggccgttgg
11941 ctttgggttgc acgactgttgc tgcgttccagg tccaaatgttgc agcccatcc tggccgggg
12001 accctggctc tgatgttgc gcccgggttgc tcccttgc tgcgttgcgtt cacctgttgc
12061 atggggccatc tactggggat gtttgggttgc ctgtgttgc cctctgttgc actaaagct
12121 aactgttgcgttgc agggccgggttgc ggttgcgttgc cccgggttgc tggccggcc
12181 ctttggccatc gtagcccttca tcctcccttca acggagcaat cccgggttgc gtttgggttgc
12241 ccaggatgttgc gcaatgttgc cggccaggagg acccgggttgc tggccggccatc agtgtatgt
12301 gaaatgttgc gaggccatccatc aggagccagg ctggccaggagg gggaggccgc cggccgttgg

12361 gacgttctcg ggaggcctgg ggaagtgagc ccaaagggcc gccccggcag atgctgaccc
12421 cagcaggaag tcaccgtccg caaggagcgc cccgctaggc tcaggcttgg cacacgacac
12481 ctcgcttcca cttcatttc ctaccttccg gatgacgaac tgcgtctcg acaactgaaa
12541 ggactgecct gagtcccaca gcccggaaaga ggtcggggtg ctgcggggcg cgccggggcg
12601 gcagaggaag ggccggggcc gagcggccac cagggtggcgc tgctggcg agaacagcgg
12661 cccggggccgg ggacgcaccc ggcccaga ctccggggcc gcaccggc tcccgggggcc
12721 ccagcggccc caagacagcc tctgggcctc accctaaata catcaattat gacctgagca
12781 aaaagttctg ttccaatctg tgctctccq aattttccatt gtttgcctg gaaggagctg
12841 aatttatctat gctttccct cttcagtca aggagacgtg ggaggaaacc gggatgggga
12901 tcagcagcct ggaggccgcc tccgctggaa ccgcggccgc cccggccagga ggctctgcct
12961 gctgggcaag ggcgcactcag ccagccccgc agggccac ggcacccgagg ctcaccaagt
13021 ccactgcagg tggcaggtga ggtcacctgg gcagctctag gccccggcga aatatctcat
13081 ctgctctgtt ctggggcaag ttaccttact tgccccacc tgcaaaatgg gcacaactgc
13141 agctccgagc ttggagctgg ggggttccg cccatgcacg ttcaaggcagg gtgcagtcag
13201 cggccacaaa tggccatgtt cctgtctct gccgtgtgt tgatactgtt ggagtataac
13261 accgtccct ggggactggc ctaccggctg gggctggc atgcggccaa ggggatgcgg
13321 cttcagcaaa gtttctcaac cgtccccatc ccagaggggc cctctgcct cagaccagcc
13381 tcctaaacag gtgcaggata ggacagcagg gaggttgggt tctgttgagc agaccacatt
13441 tctcaacctc agcactccag atattttggg ccagataatt cctgttgggt gaggcgggca
13501 ccctgttcat ggcaggatgt tgagcacacc cctgggca ccttgccaga acccagtgg
13561 accctcccccc tagtcgtgac aaccaaaagt ctctccagac actgccaaat gtccccggg
13621 tggggggagg ggtggcaaaag tcacttctgg ttgggagca ctgaccctca ggggaccagg
13681 acaaggagcc tgaagctgg gctcttgacc agccgggatc acagagtcca gcccctcc
13741 ccagagcaga gtggctccac cagcgtgtgg ccccgccgag gccgctcgcc ccctggagcc
13801 agggctctca ctgtcaaatg gggatgtca cagcgcctac tccacagggt tgctgtggg
13861 gtgagggcat aaagtccacaa aaaaaaaaaaaaaaaatggctc agacgggtcc tgcccagaa gaagctctac
13921 aaacagtca ttcacgaac gcctttatgg ttgtcgcat tcctgttg tgctggggtt
13981 tcacagaccc ctccttatgg gggaaactaa ggccccatga tgtaaatttgc ctcaccaaaag
14041 gtccccacgag tatggacaaa gccagaatga atggcaggga caagtcctca ccatccctgc
14101 tctgtgagca catgttctac cgtcccccc agctgatgc aaagtggct cccccaaggc
14161 tggaggtctc ttgttccagg cagggtttac agatcaaatc tcacacgac agagaaccac
14221 gagagctggg gcaggaggcg tgcacccat ttctctgcca gaactcccat atcggtcc
14281 cattttgttt ccagccatct tccccaaagc ctgggaagct attccagag agagccaagc
14341 tgccagaaca ggattctcaa ttgagccctt ccaggcatg agccagtgca ggcctgtgg
14401 gaggcagaag aggggcctcc tgtagaaatc tgaatgtgt gtcccccaaa tttgaaagg
14461 tggaaagctga gaaccaatgt tatgttatta ggaagtggtg tctttgagag gtcattagg
14521 catgggggtt gagtctccat aaatggatc agccacccat caaaagaagc ctgagggagc
14581 ctgttcaccc ttccatcatg cgagggcaca gctagaaggt gccccttagg aggaatgg
14641 ctttacgaca catcgttct gctggcgtcc tgctcttgg aaaaaaaaaaaaaaaatggcc
14701 ttagtaacac aattccattt gttataaatt actcgtcttta aggtatttct gttattgc
14761 cccaaacaga ctaaggcacc tcccaaggcc agtgtccagc cttatgtccct cctatgtctc
14821 tgacccagg taggtctggg tgaggaggat gatggaggat gatgtgggtg ggcacatcat
14881 ctttgcaccc ctatcaggtt ggggttggg acccagatgc tcaggaggac agggggctgt
14941 ccctggggtg aggcatgggg tgggtttccc acctcatccc ctgtccaga tccccagg
15001 ctttgcata tcctccccc accttcaactt cctggcccttc ctgcaccatg gctggggcc
15061 tctaaatctt ctaccctata ggtgggtggg tgggcacacc aagcccccac cccac
15121 tctcttctcg tgctttccac taagagactc acatgcagaa ttctcagac ctttaccc
15181 caatgtctaa aaatccacg aagcatttca cactttgtt tgactctggg tttataaca
15241 agacacttta tgccttctc tgatgtctcc caagtctctg attttccattt gctttgattt
15301 tcatcatctt agggaaatatt taaaatggaa acagttgagt agctttggta ctttctcaaa
15361 aaatcaatttggaa aaaaaggccatg tgctatgca ctcagttgt tgccatgtc ctgctcc
15421 gctgggttataa aataagatca tctttttct tctgttttt ttaagaaat
15481 ggagaaaaat gtttttttag gaaaaacttg agataactca gaggcttcataaagctaga
15541 agacaaacat gtcactcaaa caagtgcctac caattttatc cccactggag cctgggtgg
15601 cctgggggac aggcaatggg ggagcggcag gtggggagct ggtctatttc tgcacaatt
15661 tggatgtaa aacaggaacc gagtatctgg gccacttctt gcccaagtct gggcatgt
15721 ctcacatccc ccagagtcgt tccccaccgt gggggggaca gtcagcccc ctcctgact
15781 ccatcgccctt ggttttgcca tgctaaattt cttcttagaga cggggccctc attccttgac
15841 ggaaggctgt tggtttttt aagaggatc catcagaatccat cccacccacc atgtcc
15901 ctttgccttcc agggccatcc gggaaatccc tggctgtgt tctctgtgag caaaagg
15961 ctgttccctt ggacacagga gctaggccctt ttggccctt atctggctg ggaagtc
16021 atcctggctg ctccatccat ggtttccat tgcctgcac cctgtgtat tccttgcc
16081 tgctcaccac tctgtcgctg tctgtctgc tgctactgca tggacacagg gtcctctcc
16141 ctcctcggtt catgcaccc ctcggattt aacaggatc gtcggaaag ccagaggctc

16201 tgggctggcc ggatgtttt ggggcacaag gatacatctc cttccacgtt catcagacac
16261 ctcagagacc tgaaaacagg gcctccccca gtgttatgca aagtgaagtc tttgcctgg
16321 ggacacaaac tgttccagtc cacagtgaga taagtagaca aactagaaac ttttatagca
16381 attggacggg tttagtttat gtccacagaa ttaataata aaaatccaga cctagagtt
16441 ttggataacct cggttttgc cacatttgc tttcccagta attcatttat aacgtatgg
16501 acggaagtat tggctgtga tgggttgaa atagaaaaaa gaaaacagcc aggtcttgg
16561 cattgaggtg cactgtgtac tgctgtccc aggatggctt ctccccctggc caccctggga
16621 tggcagatgg ctcctccggt cccacagacg ctgtcccaa gccctgtggc ccacagcaca
16681 ttgcctaagt caggcaggc ccatgaccag aaccatctt ctagccgt gtgcctcta
16741 ccccaagct tcagcccgac ccctagcagt gcccacccct ctgtctact atttcctgca
16801 cttttgcgat aggtccaacc taaagcaacg aggactcata gcctcggtc acttcctc
16861 tcacagccca gaacccccc gagaaggcga cagcagcggg caggccggc tttctggc
16921 gtctacactc agtgcataa ctcaggagat ccccccgtat gagcggaccc catggcact
16981 ggcgtccctt acccccatca ctgcctctt tgctgtatct gggaggctt taggaggaag
17041 gcggggtagg gccaggcaga aaaaggtggg cagagtatgg ggtgagctt gctgagcccc
17101 actctggca ggcatacgaa aggagcagaa aaaaaggccc agaatgtctt aatgtccag
17161 aacattccag gaggccaatc atccctgaat tccccggatc tgacagaggg accggctgt
17221 tgggtgcctc caagcttaggc tgaggggctt gggacaggc actgggtgtt cctggtccc
17281 acacccctca caaccccgac ctgcctctt cctgcctctt cagctcgccc tccactgagc
17341 cctcggggcc ggggttaagc tgggtccatc tggagtttagc tctggaaagcg tgcgtccagc
17401 ttccgggtct gtctccat cccctcacac catcgttctt caagccccctt cccagcagga
17461 ggcggaagca agggcagct cagttcttctt tggtcttctt gtacccgg
17521 cgcagaagt gcccgggtcc caccacaggc atgagctcc ctgggttgc ggttatggat
17581 ctcgcacccc ccaggggctg ccatgttccc catgttgc tctggaaagag caggctgcac
17641 ctgacaaccc gtccctgtcc cactgccccca ttcccaactc tgccaaaagt ggaagtgggg
17701 ccccaagcatg gcccaggtaa agcagagctt ctgcagggtc atgaggggtt ggggtccca
17761 caaaggcagaa ctctgcagg ggggggttgtt gcctgtaac caggcgcacat gggcaaggag
17821 ggaggggtgtt ggcataac tttttttttt aaatgtttt tcccaagatt tccatttcc
17881 tatggcagcc tgctcctca acaggaaRcg tgagctagcg cccataggtt ggaacacaga
17941 gtggacccccc tgccccagcc aacccatcta ctccctccca gcccactgca gagatgacgt
18001 ctctttgtat tcttcattga atgcccacccg agacctttctt aatcagaaat aaaggactga
18061 taccattttt aaatgtttt tttttttttt aaatgtttt tttttttttt
18121 aatgcgtgt tttttttata aatcccagtg ctatgggagg ccagagtgt
18181 aatcccagtg tttttttata agcctggca acagagttag
18241 aatgcgtgt tttttttata aatcccagtg tttttttata
18301 aatcccagtg tttttttata agcctggca acagagttag
18361 aatcccagtg tttttttata agcctggca acagagttag
18421 aatcccagtg tttttttata agcctggca acagagttag
18481 aatcccagtg tttttttata agcctggca acagagttag
18541 aatcccagtg tttttttata agcctggca acagagttag
18601 aatcccagtg tttttttata agcctggca acagagttag
18661 aatcccagtg tttttttata agcctggca acagagttag
18721 aatcccagtg tttttttata agcctggca acagagttag
18781 aatcccagtg tttttttata agcctggca acagagttag
18841 aatcccagtg tttttttata agcctggca acagagttag
18901 aatcccagtg tttttttata agcctggca acagagttag
18961 aatcccagtg tttttttata agcctggca acagagttag
19021 aatcccagtg tttttttata agcctggca acagagttag
19081 aatcccagtg tttttttata agcctggca acagagttag
19141 aatcccagtg tttttttata agcctggca acagagttag
19201 aatcccagtg tttttttata agcctggca acagagttag
19261 aatcccagtg tttttttata agcctggca acagagttag
19321 aatcccagtg tttttttata agcctggca acagagttag
19381 aatcccagtg tttttttata agcctggca acagagttag
19441 aatcccagtg tttttttata agcctggca acagagttag
19501 aatcccagtg tttttttata agcctggca acagagttag
19561 aatcccagtg tttttttata agcctggca acagagttag
19621 aatcccagtg tttttttata agcctggca acagagttag
19681 aatcccagtg tttttttata agcctggca acagagttag
19741 aatcccagtg tttttttata agcctggca acagagttag
19801 aatcccagtg tttttttata agcctggca acagagttag
19861 aatcccagtg tttttttata agcctggca acagagttag
19921 aatcccagtg tttttttata agcctggca acagagttag
19981 aatcccagtg tttttttata agcctggca acagagttag

ggcaactatcc taatttatctt cactttacaa gatgttgga ggaagtctct tacttgcttc
20101 ggtccccacag ctgtctcgcc cattcttgcg cgttcatgtt cctttggagg aagtctcttc
20161 cttgcttggg ttccgcagct gtctcagcca ttcttgcacca ttcatgttcc ttggagggaa
20221 gtctcttccct tgcttgggtt ccgcagctgt cttagccatt cttagaccatt catgttccctt
20281 tggaggaagt ctcttccttg cttgggctcc gcagctgtct cagccattct tgaccgttca
20341 tgttccttgc gaggagaigtct cttccctgtct tgggtccgc agctgtctca gccattcttg
20401 accgttcatg ttccttggaa ggaagtcctt tccttgcctt gttccgcag ctgtctcagc
20461 cattcttgc acgttcatgtt cttttggagg aagtcttcc tttgttgggg ttccgcagct
20521 gtctcagccaa ttcttgcaccc tccatgttgc tttggagggaa gtctcttcc tttgttgggg
20581 ccgcagctgt ctcgcacatg cttgaccgtt catgttccctt tggaggatgt ctcttccttg
20641 cttgggttcc gcagctgtct cttccctgtct tgaccgttcc gtttccctaga tgggtttag
20701 aatcacttca taacatttcc aactatgaat ctggaaaagg ccaacacttc gattatacc
20761 agcttcccat gaggacagca tttgtgtctt catttgtat cactttctt ctatcattaa
20821 tcaggtacaa ttcccttcca atgggtccat tgcgttgc gactattcc aggtattctt
20881 catttttgc tactaccata atggaaattt tctccctg tttcccccac taccaggct
20941 gttgttttgc gttcacacag aagctactga tttttaaac aatggtctga tattttggca
21001 catcaccaga ctcctctctt aactctacaa gttRtaccgt ggattcttt aggtttccc
21061 gacacatgat taacagcaaa caatRagaat ttgtctctgc tttttccaa tactagttt
21121 atgtcacata caagtatcta ggattttcag aaccacagta actagttat gggtcttttt
21181 tgtttgcgtt tttccaagac tatctccagg gcagtgctcc ccagagcagg ttcaagcaggc
21241 tgccgatctg ctgagcaaga agttccctgt gagaaggtt tcaagacca cctgtactgc
21301 tcggacttcc ggaggttacac catgcacacg tccacaggaa agtctccggg aaatttcacc
21361 cccaggaatg cccattttcc tttgtttgc cagcattttc cagacttctt tgacccttgg
21421 accccctctt atgtctctgtt aacatctact catatcttca gaattcgctg tgcagaacac
21481 ttttaggaa atgtgttctt tttccgcctt tttttttttt tttttttttt tttttttttt
21541 tctgttacta gtttcttggg gatctggaaa aggaggagtg ccctgtgagg ggctaygggg
21601 ccgtcacgcc cttgcccctgc cagaggcttcc aacactggag ccctgaggccc tctctccag
21661 gctccccccta ggcagcaagg tcctcatgccc tgcaatgtct tctgtacacc ccagctccag
21721 accccgcagg gctccagacc cacatccaca tcttcttggc tccctccgtt gggatggcac
21781 caggtacact taactgcaca tctcagctga cctttccatc tttttttttt tttttttttt
21841 tctcagggt tgataccatc atctaccatc ctgagagtct tccgccttc acccacctag
21901 tgccacatcc tctcgaatgtg gcctccctgt tttttttttt tttttttttt
21961 tggtcccaagt ccagctgcca gcatccccca tttgtatggca tgactgtct cctaagccac
22021 tgctcccaacc tgcccttctt gcacacagga gccagtgatc ttcttagaaa tgacatggg
22081 atcggttccat tgcttaaagg tccccctatc ggcacggccaa acccccttgc gggagacgca
22141 ggtctgtctt gtccttgcg atcggttccat tccgacgttcc ttggcattgg gtcgtgtt
22201 cctgtccctg tgctcgcccc tgggttca gtcctccccc tccccctccc agcctatgaa
22261 ctcagagcat ctctatgtct gtaatccatc tggaccatct tcagggttgg gttggacgg
22321 cacttccccc tgggttggg gatattaaac aatctccaca tcccttagat aagccctaga
22381 ccatccagcc ctgagccctg tttttttttt tttttttttt
22441 gtccccctgag gacagagact atgcccgttcc tttttttttt
22501 agcccttcaa acagtatttgc tttttttttt
22561 actgtatgaa tagatttcc tttttttttt
22621 tgatcacggt agaagttct tttttttttt
22681 ttttacttctt attttttagaa gatccatcc tttttttttt
22741 gcatcttgg gtataagggt tttttttttt
22801 acatgagttt acagggtatt tttttttttt
22861 atgcacttgc aaaatttgg tttttttttt
22921 ttatggcata ttctccattt tttttttttt
22981 tcaatttccc cagatctcc tttttttttt
23041 gtttgactct ttctgtccaa tttttttttt
23101 gcagtttaagg aaaaaaaaaa tttttttttt
23161 atagagcttgc ctaaggagta attaataaag gaggcttatt tttttttttt
23221 gggaaagagc agacactttt caggccaaat aagatccaaa gctttgcacag cactggtaag
23281 caattaaccc tagtcaatta caaaattttt cttttttaaac aatgttaggag caacctgggg
23341 gaaaatgcaa gccgccaccc tttttttttt
23401 tctagccccca ggaccaggag tttttttttt
23461 gtcccatctt catgcacaaag tttttttttt
23521 agggacaggc tggaaaagacc cggcttgggg aaggccggaa cagggaactc cactggaggg
23581 aggcagccag ccacatggga tagtccaaag atcccccccg cacctccccg cctgtgaatg
23641 gagacttgc accggatttc tgggtctggg ccacccccc ggtctctgat ctgtgtgg
23701 caaagcccccc tggccacaaac tggggataga gggctcagcg agcctcttc tggcccttca
23761 gccagcattt actggcagag gatgcccctgc acatccctcc tccactgtatc atccctggac
23821 ctccgaagtg gccggctgca gagaqcttc caacaqcttc ttaqtgggca qatqccctg

23881 catgcttcc tttctgccag gagaagtata tttttgttagt cacattacat atactcttc
23941 tgatcttta gtttgcgcct tctctcttt tttcttagatt aaataggta gaggggttgc
24001 tttccMaag ctctcttctt cattgattct atgtgtttc tccatcttct aatttattga
24061 ttctgtctt ttctgaacta tttcttatt tttgtttcc ttaagttttt ggttttgtg
24121 ctgtgaaat tcattcaaat gtactgagtt gaaggcaggg ctaatgtgtg tcgaacatata
24181 cacccaagga agaagcttt ggaaaacatc tttgttttgc tttgttttgc tttgttttgc
24241 cacattccaa aagatttgac agtagtgctc ttgtttttat tagtgttcca aatagtctgt
24301 actgcagtc taatatttc tttcttgc tttgttaccatc ggctggagcg cagttgtgt
24361 atctcagctc actgcacact ctgcctcccg ggttcaagca attgcggccatc ctcagctcc
24421 caaatagttt ggactacagg cacatgccac cacacccggc taattttgtt atttttttgg
24481 tagagacagg gtttccaccatc gtttagctagg ctggctcaa actcttgacc tcaagtgtc
24541 taccacccctt ggccctccca agtgcgttgc ttacagggtt gaggcaccatc ggctggcccc
24601 tagtaggtt tttttttttt taatgtttc aaatgaattt gatttttttta ggtttataaaa
24661 ttctgtgcc cccaccctt cttttagca tttcatcagc gaagatggac aaaatcggtt
24721 tagctgtata aaattcacaa ttttttttta atgacccaaa acactctgtt aatttttaag
24781 tttcttatgtt tttctcgaaa agaagggttca tttttttttt tttctgttagg gcaagggaaa
24841 cagtgtata actactatt tagtttact ctttgcgttca ttcaaaaattt ctatgccta
24901 tttagcatcta atctgtgaaa ggctttaaa agtgtattttag agactctcat agtttaggtt
24961 ttgccaattt cccctttgtt ccaaggattt gttttggat gttttgtgtt ctcttgggg
25021 acacgaaggc ctgcaactgg catccgcctt gtgggtgcct tcggctactg tggaaatgagg
25081 tgcttgggtgg gtgttagggaa gtggcccaaa atctacttag ttttcttggg atcccttctg
25141 ccacccggagt ctttgcgtcat cattttaca tttttttttt tttttttttt atgcagttt
25201 attttatattt ttgagactga gattttttttt tttttttttt tttttttttt tttttttttt
25261 gatgtattttt atccctattt gatttattt tttttttttt tttttttttt tttttttttt
25321 tctgtgttgc gggtaagaag cctgactctt ctttgcgttgc tttttttttt tttttttttt
25381 gggcatgtca ctaagtca ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc
25441 cagtgtgcac cccctccatc ggccctgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc
25501 ccacacggct ggtacaggca ctttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc
25561 tcctctgtct gcacagcatc ccgtgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc
25621 acccatgcac accggccagga aaggacgtcc ttttgcgttgc ttttgcgttgc ttttgcgttgc
25681 gcaagagagc ccagctggga ctttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc
25741 acgggggcag aagtccaaagg ggggttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc
25801 cgagccttgc tgactcttgc gggactgccc ttttgcgttgc ttttgcgttgc ttttgcgttgc
25861 gaaacaactc atccatgagc acattttca gataaaaaacc aaccaatcca gcccacatgt
25921 ccccaacccatc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc
25981 cccagaaaaac taactcgatc ctttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc
26041 gccaacccatc agccagctca ctttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc
26101 tcaggctgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc
26161 gggccccgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc
26221 gtgaacaata gtaatggcca cattttaaaaa ttttgcgttgc ttttgcgttgc ttttgcgttgc
26281 ttttataatgggaa acctcaggggg aggccaggggg ttttgcgttgc ttttgcgttgc ttttgcgttgc
26341 tctccaagga ggcccatgca ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc
26401 cccgctggcc acagtgcact ctttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc
26461 acgttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc
26521 ggaggcatca gcccacccatc ctttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc
26581 ccccccacccatc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc
26641 caccttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc
26701 ctttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc
26761 cacctggggc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc
26821 ccagctgaaa cacaacccatc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc
26881 ggcagggttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc
26941 attacctacc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc
27001 atgttttaac gtttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc
27061 ctgttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc
27121 tacacgtgtt ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc
27181 agcagatata aaaggggaccc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc
27241 cacagccatc aaggccatc agaaacatag ttttgcgttgc ttttgcgttgc ttttgcgttgc
27301 gggcttccatc ctttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc
27361 cccttagcgca gcccacccatc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc
27421 cgaggccatc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc
27481 catcaagccatc aagggtgtttt ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc
27541 cacctgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc
27601 cagggatacc cgctgaaagg ctgttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc
27661 ctggctgtgtt gtttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc ttttgcgttgc

27721 cttccccggag cctgctgcc a cgtcaggcca gcacacactc cctgcaactcc aatctgtaga
27781 gcctgcattt gaggcaacagt gattgttacc tcaaaaacagt gggaggcac ctcaggaaac
27841 cYaggggttc acaaggagaa tgacatcatg atggatgggt ggacaggcat acagggcggt
27901 caagacaagg atccccgggc cccgtggcat cgggtgttgc aacaaagatg gcaccgatcg
27961 tggggcagag ttcatgacat gtggaccctg ctgagggtac catccagccc aaacgcagct
28021 cctcttggca actgctgccc agcaaatacta catctgattt taaatggct ccccaacaact
28081 cccacccctc tttcctatag gcaagggaca actcatcaga cccacgggag cgtggcttgt
28141 caccgtggct cccggctgtt ggcgtctgg gagggtgggt cagaaccccc ttgtaggcgt
28201 gagggtttac ctctagaccc ccagtgtaccc atctcgcccc tggctgtac aagatggagc
28261 tccgaggagg gcacccaggc cccaggaaag ccttgccatt tggtcatcgc cctccgtttc
28321 ctctgtaaaa tgaggatgc agtccccca tcacaaagct gttgtggga taaaaccgYc
28381 atgcagggtgc aatggaatgg tgcttgc aacaatgaaa cggttttca ttgctggtaa
28441 accatagtgc cctgcagctc cctgtgtgggt cctgggggtt gggatgtcac ccctaaacca
28501 cagaagtctc tgcccagtt tgagggttcg actagattgt ttagtccacg ctggcacagt
28561 agccttaaaa gccagttttt cttaaagagag aagggttttgc tggatgttc cacagtattt
28621 ctttcttat caccacaaag cttggccagc ttggctcatg gagtctgacc gagggaaagg
28681 cccactgcag ctgaaggggag cagaatagga gagaggagaa aggccaggga tgaggggca
28741 ggcttctctt catcctcaac cccaaacctca caggtcatac agaacagtcc tgacatctga
28801 gaaggaagga aatgcaaaag agcttgatcc tggctgctgc ctggacactg agaggggaca
28861 ggttcttagc agtagggatg gcttccaggg cccaaactgtt gggccagctt gtcaggcct
28921 gagagggtgg cccagctccc agctgtgtt agcagatgcc agggccccct gcacccca
28981 gacacatggg caagatagtt gggactggc ttgcagggtt acctggctt aaaaaaaaaa
29041 accaggcttc agggcaagcc caggggggca ctggccttaca ggagaggagg gcaggccctgc
29101 ttgcaagggtt actgcagagg ggactcatcc ctactgcccc ccaccaggcc tctgcctgca
29161 aatggggctt ttgggaggtt attaggggtt gatgagttca tgggggtggg accctcatga
29221 tagggtagt gtcctatagg aacagacacc agagggctc catctcttct ctctggccgt
29281 ggaggacca gcaagaaggc ggctgttgc aaaccaggaa gagaatcttc actggggat
29341 gaccctgctg gcaacatgtt ctggaaattt ccagtcttca gaactgtggg aatgaattc
29401 ctgtgattt aacccttggg tctatggcat ttgttatac cagcccaagc ggaatgagac
29461 aaaggcttta tggggaaagag actgagccac ctcccacagg ccagactgca gccttccttc
29521 tatccctgggg ctgagtgtcc gagatgggg cagcaccaga tataggaaacg aaaaaaaaaa
29581 ctggccaggc gcagtggctc acRccctgcaa tcccagact ttggggggcc aaggccgttg
29641 atcacctgaa gtcagaagtt caagaccagc ctggcaaca cggtaaaacc ctctctctac
29701 taaaaataca aaattagcca ggcattgtgg cgagcgctt taatcccagc tgctcaggag
29761 gctgaggcag gagaatcaact tgaacccagg aggccggaggt tgcaactgagc cgagatcgc
29821 tcattgtact ccagccctggg agacagagca agactctgtc taaaaaaaaaaaaaaa
29881 aaagaatgaa agaaaatgttc attttttttt acgttttgc ctaagcgtgt aatgttgc
29941 ctaccgtcc acacagaccc tcagcacacg cacggggccc tcacaggctc tcaatgggac
30001 ttctcaaaact aaactgaccc tggctcttgc tgcaagccac cacaccctgg gaccaggaa
30061 ggcttggcg agcctgaccc ggttctcttct tccccactgg gctgaggctg gcacccagca
30121 gacccagcag aaccactgccc tcaccatgaa ttgtggctt gatgatca gagccagccc
30181 tcctgcccagg aggcggagag ggacactcag atgctacgag accagctcca ggtacccccc
30241 cctcgaggga aaggcaggga accagaggat atattagttt gatatctccc actggggct
30301 ggaaggacgt ggcattgtcat gagacctgag aaacagaaca ggttagctca gcaaagcagg
30361 gatggcgctc ctggcaggagg aggccaggccc tgcaaaacctc ccagggttgg aagcagcgg
30421 gcaccccttca agcacagagg ggatgtggaa gggatgcccc agcagaggag gtgggggggg
30481 agcaactgagg ccagaccacg aatgtcaggag aaggatttca aatctcttct aagagcagtg
30541 ggaggacatg gagaaggat gatggggggg catttaagaa gtttacttca ggtgccttac
30601 gggaaatgaa tggaaagggtt caagatgtgg gatgtgggg ttcaatgggg gggatgatgt
30661 cagagaagggg acggatgtga gggatgtgtt gatgtggaaat aggggttcatc tgccccgct
30721 gagcaggggag aggctggggc tctggaaaggg cattccggca gttcccacag ggccggaggga
30781 ggagatgcgg ggctggccct ctccctggctc ctggccaggc ctttttttttcc accacgc
30841 atcttccac ttccttagcc ctaccctgg ggccttgc ggccttgcctc atctcaagaa
30901 ggaatcttgc ttccttgc acatttgc gatctaattcc tgccttgc ttttttttcc
30961 gaggtgacca accactccaa aaggccctgc gctggggcatg gatcttgc gatagaccc
31021 tcagcctcac attctgtttt gagggtgggtt cgaggcagct gacccctggg gtgtcaggct
31081 gctaggcgtg ggcccttggc tggggagctc cactgtggc tgccggccag cagccagctc
31141 ccaccccttgc cgttgggtgc cctggaccctg gcctctggaa gcctcagatg gtccggagat
31201 taactgggc tccctggctg ctgaggagat aSagaggcga tggctggaaa tgctccca
31261 agtagccag ggagaaggag gccaggccgc agccaaggcc ccctcagccc tccctccctg
31321 gtaccccttca actgtgggtc gcttcccttgc atccctcccc caggaaacct ctgtccctc
31381 ctcccttca caccagcaca gcacatgtt cccaggccat gcccactctc tggggaggcag
31441 gcctgagccaa ggccgtgggtt gggcttgcga tggcttgc gtcctctgg gtggctgt
31501 ggcYtggat cctgctccgg agccctcaac ttggccccctt cccaggccag gtcctcagcag

31561 cccctcccag cagtcaggaa ccccatgggg cactgtggcc cagagaccac tgaccactgg
31621 cctgggtgg ttcctgtccc ctgcgtgtgc tatggcaaag tccaaccac agtccggggc
31681 cgggctcctt ctgctgtcca gcctctcagt ccccgttagt taatatgata catcaggaac
31741 agacatgggg ggcaggcagg gtttgtgggg catgtgagag gcaggggtgc tccacagtgt
31801 cctgtcctc cgtgggcctg aggagtgagc cccgtcctgc cccctctacc atcccagggc
31861 tctacctgtc tttcaggagg ctgcaaaaattt gccaaaattt cctgcccage cctggccaga
31921 ggcctgcagg gcccacccgg gcatggaaac cactgggat aatcgtgta attcccgaca
31981 aaacactggc gtccctca gggacacagg gcaacattt ctgcagagca gagagtgcac
32041 tgatgtggca aaaagagcaa ctgaagacag aaggcgacct ccacatgcct gcctcctgca
32101 catccccctc ctccccccggc cccacccact ttctctgcca gacaggagac aggcggtagg
32161 gctcccccaac ggtccccctc gtgtcctccc cccccccca cccacacaca ccattctca
32221 ctacgatcac caagcggaaat agatctcacc ttcccttcc actgtctaaa accctcaacg
32281 gctctccagt gccttcggaa gcagggccca gccccctcgc tggcattcgg agccctctcc
32341 acgctggccc tggctggctt ctccatgcgt cacttcagat acctcccccac atggcaccct
32401 tgctcctgac agcccagcca gccttcccca gagccggctg ccctctcctg ctgcctgtgc
32461 tttggttcag gcctggaaatc ctcaacttcc ctgtgcagcc cctcaatttct ttagatatct
32521 caggagcatt tgaaaaaaa gactcatctt aaaaaatata tctcaaaggaa aagaggcagt
32581 gaSctccttg cttaaaaccct tcctacccca aggcatgaga ccgcagtccc tggcgtgacc
32641 tccaggccag gagggtctgg ccctgcttag tcttctcggg ctgcctcggc cctgctccct
32701 gtgctccagc ggctagcccc gctgctcctt gcacaggcag cctgtctcca ctcaggaac
32761 tttggacttg ccgtgcaact gctcccgagca gtgyctagcc catagttagt gctcagtaaa
32821 catttgctga atgaatggaaR tgaatactt tgaagatata gataaacagg aaatgggcaa
32881 aatgtatgac ccagcgtagg cctgtgaccc ctctcatggg gaggccccca gctaaggccc
32941 tccccctcccc tccccccagc agccctgggg gcctcctccg catccaccct acccaaggcc
33001 tgcgtgtgct ggggagaagc agctctcctt tcccaggtt cagRgggagg ggaagcccg
33061 ggacaatgtc cttgagttact gtctgagaaa ttgatcctgg aatctttaaa cttgggggg
33121 aaatgttcaa agccgcagcc atctctgtgM ggagagaggc caagggtctt ctgtttttggg
33181 gaaacacatt gtaatgcgc ttttgcttat gcaaagaYgc agccaaagttt ctgttctgga
33241 atagacaatg aggaaggagc ggctctgcat ggcaatgtaa agttgcctc acttggtggca
33301 gagttccacc caggggatgg tgagccctgg gaggtggct gcggctgctg ggcagagggt
33361 actttgtcct ctgccaagggt gggggagggg tgcaaaaaagg gggaaaccgg tgcaaggctcc
33421 ctactgggaa ggcctactgg cctgccagcc acagaggaga caggcagggg ccctggaaaca
33481 gcagccaccc caaccttcag gaaagttccc atRtccttcg ctcagctgag ctggggcca
33541 ccagaggcag cctgtggccc acctggcaaa accatgtcca ggagaagcc aggctgaaa
33601 agtctgtcag aggccctgggg gaggggggcc tgctgctct tggcttctct gggacccct
33661 gaaataagtg aggagaaggc ttttctctt gtgccttcgc ggagccagg ctggaccac
33721 caggatggag gcagagaaga gagtgcagcc aggagcaggg gacagaggc tgctggccct
33781 cgtggaggaa gagtttaggc attcatgtt gcctaaacgg ggctgcccaga gcctgggctt
33841 atcatgagaa tcaacgttcc cattgcagat ctttttcacc ctcgctctt cagcagatgc
33901 ttcagagggg atgctgaaag aaacgtgact cttgctcaaa atatcaaact accatgaANc
33961 gggaaatggcc aggtacgaa ttacttaggt cagttcatac ataaaaaaagt gaggtggta
34021 gagtccccctc cttcaatacc cagcgcagga acactctgcc caagccacc ccctcactcc
34081 tcaggcgcacg tctctgtgc ctgtccacac tctgctgata cctctgcaga tgcctcagg
34141 tggggacggg gctcctccac ttccaggcca gctcatcaat gcacagagcc cagcagggtc
34201 tcaagagatg tgctgcaccc ctcagtgtgg cactccagca ggcacgttga gaagacaggt
34261 cgcattacaa accttcccg gttgtacta acgacggca gaggctgccc ctccccctt
34321 cctgaatttg gggggtagc agagtgtggg ggggatattt ccactctgag gtcagatttc
34381 atgggtgcca gcatgggctc tgccacaatc ctgctgctg acttggggcc agttacttaa
34441 cctgtctggg cctcagttt ttatctata aaatggagat aataaYagtc cctgtgtctc
34501 aggggtgtt caggttgaag caacaatgaa aattatcagc tgcgtccatg tgaagagatc
34561 accaaacagg ctttgtgtga gcaataaagc ttttaatc cctgggtgca ggtgggctga
34621 gtccaaaaag agatcagca aagggtgtgg ggactaccat tgcgtgttat aggtttggga
34681 tagacgggtgg agtttaggagc aattttttt ttttttgag acggagtctc tctctgttgc
34741 ccaggttggg gtgcagtgcc atgatctcag ctcactgca gctccaccc ccaggttcat
34801 gccattctct tgcctcagcc tcccgagtag ctggggctac aggcccgc cagtgccct
34861 gttttttttt tttttttttt tttttttttt agtagaggca gggtttcaca atgttagcca
34921 ggtatggctt gatctctgtca ttccatgatc tgccgcctc ggcctccaa agtgctggga
34981 ttacaggcgt gagccaccgc gcccggccag gagcaattt ttgtgggctg gggacggatt
35041 ttacaaaagta cattctcaag ggccgaagaa tattacaaa tatcttctta aggttaggggg
35101 ggacaatatt acaaagtatc ttcttaagga tgggggtggg gaagaatatt actaagtatc
35161 ttcttaaggt ggggggagag aatattacaa agtacatctt caaggggtggg gaggggtgtat
35221 catacaaagt agattcaca gggcgggtcc ggcggggcgg ggtggcaata tcacaaagta
35281 cattacccca agggcggaga gggtgtattt ttcacaaaattt caattgattt atcagctagg
35341 gtggggcagg aacagatcac aatggtgaa tgccatcagt taaggcagga actatctatt

35401 ttcaatttctt ttgtggatct tcagttgttt caggccatct ggatgtatat gtgcagggtca
35461 caggggatat gatggcttag cttgggctca gaggcctgac attcctgtct tcttatatta
35521 ataagaaaaaa caaatatgaaa tagtgggtaa gtgtgggggt ggtaaaaatt ttgggggggtg
35581 atatggagag ataatgggtt atgttcctca gggctgctc gagcaggatt agggggcagca
35641 tgaaaaaccta gagtgggaga gattaagctg aaggaagatt ttgggtaac aggtgatatt
35701 atgggggttgt tagaaggagc atttgcgtt tagaatgatt ggtgatggcc tggatgcagt
35761 ttggatgaa ttgagaaact aaacggaaaga cacagcgccc gaataaaagg agaaaaaacag
35821 gtattaaagg actaagaatt gggaggacc accgacatcca attagagagt gcccaagggg
35881 gttcagcata attatttgct tggggcga gttttggc tctatcctt agtttatgtt
35941 gtcatacacc aggccagact gatttagta aaaacaacac cttagtgatt caaaagctt
36001 attgctcaca caaaggctgg tgggtgtctc ttcacacggg catgcgtgac aaaagtgaaa
36061 tccccagcac ctagtacagg ctgaatgaaa ctgtcatcaa attacaggag gtcacacag
36121 cccaaagctgt ttccctggga cccatgcac agaatgggtc acctgggtt cgcccttaca
36181 cccaccctgg ctctccctt ggcacatacc atcgggagag gtctgagtga acagcttgc
36241 catgccccctg tggtttccag agctggggg gctgctgtc ccagcacRt ccgcgagtgt
36301 gaccaggccc tggagcttc cccctggct gcgtcgccct tgggaaaaca acagccatgc
36361 caccacttag ggactgaccc tcccacccgc tgaccggat cctcatcaga aaggaaaatg
36421 ctgcagaacg gggaggacc agtgcgtctg aggctgcagg acgcagaagg acacggccga
36481 ggccctcaggc cagagcccg tgccccccag gggccctctcg gccactgcga gccactgaca
36541 gggaaaggcc gaaaaagtc agggattcttR attcttgctt gcagtggctg caggacggcc
36601 ctgcctccct ggaactctgc tgagcttgc tgctgtccct ggccacagag ccctgggggg
36661 ctcaagtggg agtttccctt caaagagggg ccctaagtgg ccaggctcg tgccacccccc
36721 ggccctggca tccccccaca cgcggctctc agaggctgtg tggcccgcc catccctgt
36781 ctgagctcggt tcacctgggt ccgtttgagc accaccatgt aatttgagaa caatgcatt
36841 ttattcccc acacatata ttctggcgt gctctgtc acgtgtgacc taaggaaaga
36901 gggaggaaa atgcgttggg gggaggagag agctgtggca tcaggggaca ggaacgggt
36961 gggggagggg gcagcaagca gtcgggggg tgggggggg gctgtggc ccacaccagg
37021 ctggagccg gaggcctatc cttaccgtcg ccccccgcctt agggtagat gcagatagtc
37081 cccgggggttc tgggctccac ttcccactga gatgagaccc tctgtgagg gtggaggcca
37141 acccacatgg tggcaagca ggtggtaga atggagaggg atctcatgtt aatgaatgg
37201 cctcaggcag cacatggatg cccctcgctc cagaaggggg tggggggat gcagcagctc
37261 ctgctggcga gggccaggcc gtgggtagct ggcactggct ggcagaggct gtagagtgg
37321 caccatcaa ggcatagggc ttggactggg caggggtcca gcacccctaa gtggcccagg
37381 cgtcagaccc gaccacccca gacaaagct gaggaggact caggccagcc tccctactcc
37441 ccaggcatcc atgcgtatg aggaagaagg gagaggccct ggcagagcag gtagaaaccc
37501 ctcagagact gagaaggcat gacagaggcc tgcaggccat cacccgtc tggcccccctc
37561 gacctggca gggggcccca aggactcg agcgcaggcc aagagtaagg tgaaggctgc
37621 aagaagggtc ctgcattaca agtcacccgc aagggtcgct tgctggcage gtatccctcc
37681 ctgaaagcgc agtcatgcag ggtgtttcc aacacccaga gcagatgttc cagccctct
37741 tgggcttggc gggaaacggc ttggccctt tgcggccat gtcacaatcc ccctgcacca
37801 ccagagtctg tttatgtggc ttagctccct tgccagactc tggcagatgg aaagatagat
37861 gaaatccctgc ccaccgtc acaggaaacc cagaacccaa ggaagagggg ccgaattatt
37921 ggaagcacaa gggtccgtga tcacttctgc tcaggatcg aaggccage atggcagcc
37981 cccaggccag gtcggctgcc aggcgcgggt gggaggctt gcagctgca gccaggttc
38041 ctccctctg tctggctgca gggaaagttt cttgtctcg aagttccctc actctgtgaa
38101 cattcccaac gggcacccca tgcggctcg ggtgtgtcc catccgttta acttccctg
38161 accatgtga tctccatcg cttcggtt tgatgtgt gtcggctcc actaggccagg
38221 gactgttctg cattttgt gatgcawcc ggtgtctcg tactccctg tttccccagg
38281 cttcaggacc aaggccaaac tcagccgcac gcgccttctc ccacccagg gggctccca
38341 ggtggccgag ggctggcacc tctttgcac aaggggctcc tctgtgtgg atgtgcct
38401 cccactttcc acttgggaga caccacactt gtgacccaaa ccagctcaag gatcacccca
38461 ctgtgactat ctccacKacc aggtggagac agtcgaccgt ctcccttggc tgccaagRcc
38521 ctgtcctcRc tgcggcaagg caagcagggt gccccatga gaaaaaagct tgcggccagg
38581 gggagctggc tggggctt cccagacatc atggggcag tgaacaactg ctgcctcgc
38641 tggcggctc tgatcttctc cgccatYgc gactacttcc ccgcgtgcga tatcttccc
38701 ttccccacacc ctgcgagatgaa gtgttatccc cgctgtgcca gtgaggaagc cgaggccac
38761 agagattgag aggctcacat aagtggcat gggagtcact acaggaggctc tgaggagca
38821 ttggcaaaac ctgctgcaac tagaaactgt atccccggac ctggccattc agacgtc
38881 atggacgtg gagacacagc ctctgtcata gaagcgaagg cataaggaca cccagccac
38941 aggagttcca caccacacca tagatgtcat tttatgggaa gtctgtgaga tcacagtgc
39001 tttatggaa ttcatatgt cacttattca ttcaaaagag atttaactgag cacctactac
39061 gtgcaggccc tttccaggacttactgtgaa aaatcagggt ggcagaccct ctctgggtcc
39121 gataaaagt caccttggc ggagtcctt accctgtcata gacacaccac caaggagacc
39181 cctggactgg ggaggaagac ccaggattgc ttgactgca ctcagctcca gggcagcc

39241 gggaggggcc tcagggagtg agagccccacg ccaggtcagc ccgacagccc ggctgggcac
39301 atcatcgct ctgttcttgt ggttcagctg tcaccctcett gctggagcct gcacaagaca
39361 agctccgagg ctgcctcaag ctgtcatccg ggtgtgtgga tggctggccc tcctcgcca
39421 cagggactct aactccatt ctgacaacca tcagttcccc gggggaggccc aaggacagga
39481 tgctgggagg cttccacatc agctatcagg aggtctgcgg aggagggagg agggaggagg
39541 tggggagatg tccactgtgt gggagggtgc tgcacatttc aggaggaagc agaggtgaga
39601 tccaccacca agacccctgt tcgagtgttc cccacctgtc caccatcaact ggtctgtcg
39661 cctcccccagc cctgtgtgaa catgtgtcc ctgcggctgt gatcccgtgc tcaccacgc
39721 atccccagca ccttaggctg ctctatccgg gtgcaactag ggagttagt ggcaggcggg
39781 Ygggccttag aacaggagga cacagtagat ctatcgaaaa gccaaggcaa agcaagaggc
39841 agagccgtg gcctgYgggg agaggcgggt gtggccaggg gactccaaat gctgtgggg
39901 ggtcagttag gcctgagatg agaaccYgcg gagcatgcta ggggatttg gagctctcg
39961 tcttcagggc agacagtagg ggaggaccag cacaggggcc accttcccac aggtatgtga
40021 ctccatctgc cgtgtctgg ccatgtctct ctcagtcata gtgacctgat gcagggcaagg
40081 gaacctcaca gggccactgg ggcaaggtag tgaggtcaca ggtgtgcctc cttgcacagg
40141 gccagcacgt gcaaaaaccag caaatgttag ttccctcct caggtgggg ggggtgtttc
40201 caaggataga aagataaagc cccaaacctc ctaagagaga ggagtctatc ttttagacc
40261 acaaataaac ctcgtcccg ccgtccaaagc cccacaagct ccagcaggcc ccaatctcg
40321 cactgcgatt cacatgcggg ccaggcctt ggcgcgcgtg tgccaccacc aggagtctc
40381 attagaggca actgcgggtgg cccttgcctc aggagccttc octgaccaggc cgccccaggc
40441 acccttccct gatgaaccaa ggtgtaatca cgtgtctgt tcccgctcc catactctcg
40501 tgctcYgagc tccccagtg cagggatgtcg gaggccagga gtccaggag tgccgaacat
40561 gactgaggagc agctccaaat cccagcatgc tcctccaca cccgtctct actgagact
40621 ttccctgggt tcaactcgag tgggtctcc aacagccccg ggagtggca ctgaccagca
40681 cagggcgcag caggacaggc agtgcctcc ttccctctag gctggaaaga acactcaggc
40741 caccacacca ccagacttgt cctgcagccg tggtctgett cacctgtgca gtgtgtcttc
40801 ttaatttcaa tctgtggccc acattgaaaa ataatttagat ttcccaaaaaa aagtagattt
40861 ccagcttctc atgaaagagc gagctggcag cagttgtctg cttattccta cagaaacaca
40921 agcgagtggc accacgagga gctctgcagg ggggctggc ctctccggc gcccagttc
40981 ccaccactc caacagcctt cagccacctg gggccctcg gctggctac tggagcagc
41041 tgagggcaact gccccttct ggccttgcg tgccgcctcg cacacaccac tggccctcat
41101 ggctgtgtct cttgggtggca cacagggcca ggctgttaagg ccagcgggac aagctgtagg
41161 aacagaaaaca gtgcgcagca cccccccagg gccacagcac cactagcaga ccctgtaccc
41221 cgcttgcctc ccccaactta tcttcccaag gaaacagagc agggaccgt gttctgtc cagtttcaga
41281 ctggctgggt tgcattctcg cagcactcg tgaactatcg tttggggcga cgcctccgc
41341 actcctgtcc gaccactgccc cggccaggtt acagatttc caccacgacc agctgaaaac
41401 tggtgagaaa ttgagcccat atgccttcgg attgttgca aaccctggca aggctctgg
41461 aagtccagggc caatttctgg attaaattag aaatgaaaaa agggaaaaata tctgaatct
41521 ggcatacgaga gaaatcatc attcaaaaaa aacggtcgta gaacatactg tacagagaat
41581 gttctgtgc tggaggggt caaagggttat agccgcact gaaggcgcag aagatcttg
41641 gattccacgt tctccatagg ccttaggagcg tgcattccct cagcagaaga gatgtttgg
41701 gggctcaggc atttcccaag aacataagca gttAACACAG ctggggagaga ctattctgca
41761 tgcccgaaaa cctgtccctc ctgctctgc gttctacagc gtttaagaSc agttoattct
41821 atccctttaaa aagaatcccc gaagtaggcct caccggctc ttgagctgat aaggctagtg
41881 tgggggtggg ctgggtctg agcagccac tgcgtctcgt acttccctaa tacgtacacc
41941 cacgtgcaact gatgtgtga aaggcacatc ctcactttgg aggcctggg ccatacacgc
42001 tcaaggcagg cactaagcaa gtacatttc gcacacgtc gtacaaaaaa ttctttagc
42061 acctcatagt gaatgttctc agctttgtgg gccatatggt ttctgttgc actaccac
42121 ccactgtgt agcacaaaaa tggccacagg caattcataa atcaataggc gtggctgtgt
42181 tccaataaaa ctttattttac aaaaacaggt gaagggccag gtggccctgg ggccagagtt
42241 tgccaacaca gctttagaat ctccagttta ctttgacaaa ttcacgtgaa ttttgccttc
42301 atgcaagtaa tacgtccaaa tttactttaa tgaaaactg atgaaaaaa atgcttacca
42361 tccaaaaaag gggccacttc agggggagat gYcttacatc tgcgtgttgg cctctggcag
42421 gcctgagagt acgtgttctc ggttagcaata ggtttagcgc tggccttgc gcaggctgg
42481 tatcaccac tgcaatggtc tgcctccgt ctacaaggag tgcgggtggc ccaggggctc
42541 acaaagacag ttcaagggtct ttgttagatcc tgatatttaa gcaacacatt gaggccttag
42601 gcttctcggt aagagttaga ttgaaattag cgctttagt aatgtcttga agccacaaaag
42661 gaagcttcag cttggggccc aagctttata ttcttgcag tcttgcaca gatagatttt
42721 acttttaaaa atgtatctgt agccaggcgt ggtgtctgt gcctgtatc ccagacttc
42781 aggaagctga gacacgagga tcaacttgc ccaacttgc ggtttagt gactccacca tgcagcatag
42841 catgaacctg tctctaaaac aatctttaa aaaatgagtc aggcatggt gtgtgtgcct
42901 gtagtcccag ctgttggga ggctgaagtt tgaggtcgca gtgagctatg attgcagcac
42961 ttagtcccg cttgggtgac tgagttagac cctgtctcaa aaaaqaaaaa aatqcatata

43081 taaacatatt acttccacaa cggctctgat cttccaaggc attcatccct gcctcccatt
43141 tacctgagcc tgcagaagag tccctgtctg tacctgggga cgaacttga ggtgccacac
43201 acatctgagc tgccctcttag catggcaccc tcgcctgtca ggcccaagcc cactcttttag
43261 gggccccag gcccacacat atggccacag catctggtt gtgagcagag actagccccag
43321 gcgcgtggcag agcaagttt taaaaatgtg agaaaataccaa aagccccagg gtcagtatca
43381 ggcattgtca gcaatgtcac ctgtgtccca gcctggcagc acccgtact tcagggagta
43441 tggttcccaa gcagggaaagg tcaaacatga cagtggtctc ccaatgcctt cggggggcgaaa
43501 gggcgccct gggcagagga gagcccccta tcctcccgcc tggggccagg ggaggcagca
43561 acgcaagggtg gaggctgcag gctgagccag agtagaaagc tgggaggagc aggagcagaa
43621 tgaaggcaga gggaaatagaa gtcactctgc cagtgcgggg gctggggggc cagcaaagct
43681 ggtatggccca cacatctgccc cacggctctc actccatctt cctagcctca cagtgttgg
43741 aaggggagacc cctctgcct tagacggggg aatgtggca tccccaaaggg gcctcagcat
43801 ctctgaccag gctaaaggcc gctgagccctc aggagaccag actcatccca ggggttgcag
43861 cagccaggag ggctacagct gctgtgtctc aaactccagt gtggccggg agcagtggct
43921 catgcctgtg atcccagcac ttggggaggg tgaggcggtt ggatcacctg aggtcaagag
43981 ttctagacca gcctggccaa catggtaaaa ccccgctctt actaaaaata taaaaatttag
44041 ccaggcatgg cggcacgtgc cagtagtcaaa aggtacttgg gaggctgagg caggagaatt
44101 gcttgaacct gagagggtggc ggttgcgttg agccgagatc atgcatttc actccagccc
44161 gggcgcacaga gcaagactct atctcaagga aaaaaaaaaaaa aaaacaaaac tcactgcaga
44221 gagcctctgt ggacaaggcc cccgccttgc cctgcccacag gctcagcaca caccctgtc
44281 cctcttagca cggctggctt cgcctggcc tcagagggtt gtggacatgc ccaagcagag
44341 gctggggaaa ggccacagag aacggggatc tgagggtggca tccagggggc acatgcaggg
44401 cagcagggtt tccctcaggc tatggggctc gagcgcaggg ctcaggtttc tgggctcggc
44461 cttagggacag gctgacttca tagatgacca gggcctatcc tcagaaggcc cgcaccccttgg
44521 ggcttgaccc tctgcagtc tgccttgcgg aatgtggca ttccttcgtca ttcaccctt gcatgtgtt
44581 ttgttaagtg aaggctccgt gggcaatgca gcatggtccc atctcccccag gacaggctt
44641 cagctgcccc cagcaggggc aggttgcggc atgtattttcc cacggcatct cagggcaggc
44701 gtgggtggc ccatccggcc cacggctggc agggaaagggg catgtctgtt ggatggctgg
44761 ccagagcccg tggggctggca gaggtgaRcc tccctgtgtt gacacctgca tagagatgg
44821 atggccgtg gccatagct ctgggctggc gcaatgcacc agtaggaagg gggcgctgc
44881 ggctctactt ccccaagacgc cttggccagg gcacagccaa tgggttcttc caggggtggc
44941 aggaagctct cagaccccaag tgggtgtgcc cttcctgtc ggggtgcact cactctgtca
45001 ggaccccgta cctgaggagc ctccctcg caggattcca ctctgtgaaa tcctctcttc
45061 ttcccccag cctcccttag tctggctgtc cctgtcataa acttctggc acgctgagcc
45121 acccatcagg gatcagccac aaaatacaca ttgttttaat ttcaatgtacc ccacataga
45181 gctaaatgtt ctaatatttgc cattaaaat agccatcgca caatttaaag atgaataaga
45241 aattcatgtt aataatttcaaa ggttccaatt ttgttcaact tagacaata taaaaaagca
45301 aataaaaaaaa acagcgatg gagagacaga gagacacaga ctaaggaaga aaggaaatgg
45361 ctaaatatgc cttttgtatg gcaccccttc ctgtcttttgc aacaacgcattt cccacattt
45421 cattgtgcgt tggggctggc gaggttactac tggctgtcc tttaggccag tggggccac
45481 tggtaatatc ctgacgatgtt tgcacgcgc cagatatttgc gggggaaataa gaatgcattt
45541 catgtatatac attggcttctt ttgaaatttgc tatctcccccgc cccacacccctt ctttccttt
45601 tttttttttt ttttttttgc gagagaaatct cactctgtcg cccagactgg agtgcagtgg
45661 cgtcacctca gtcacccgc acctccgcct cccaggttca agccatttcc ctgcctcagc
45721 ctcccgagta gtcgggatca caggcacgtg ccaccacact cggcttatttt ttgttatttt
45781 agtagagatg aggtttgtc atgttggccca ggctgtctc gaactcttgc cctcaagtga
45841 tcgcggccgc ttggcctccc aaagtccctag gattacaggt gtgagccact gcactaaggc
45901 tatctttctt cttactttgg gagaaattca ggtggggc tgctatggct aaacccaaga
45961 acagtccgcg gttgggtggg aacttggggatc tggggatgtc agcttgcacag tagggcagaa
46021 actcagaaga ttaacttctg gggggggcag gaggcatcMa acacatcacc caagggagaa
46081 gctttccccc ttgggggggaa aagactgcag gatctttccctt tgggttcaca cacacaaaa
46141 tggtaggacat caaaaactat cccaggccgaa ataaaaacaaa aggtcaacaa caggactgg
46201 gggaaatattt cccacacatg tgacagacaa aaagtttact tacttaagta cacagactt
46261 tcagatgaat aaaacacccca caaatcccg cagaatcaag ggcacatgaaata ggcacccaa
46321 aaaaataaaaa taatctctgt ttacaaatatac acgtatattt ttaaatgcag gtcattaa
46381 atcaaggaaa taacaactta caacaaaaaa acttatctcc gttgtatataa aattaccact
46441 gctcaaaagg actcagttgtt gtgaggaggt gggggagacag tcatttgc taatgtacgg
46501 gaatagatga aaccataaca gagaacatcc cagtgtatgg agctaaaaaa atgggcata
46561 ctgttggccca gagacttcgc ttgcacgttca ctagcttgc gaaataatgtc agcacggcgc
46621 atagatttcaatctatgtac gttgatcatgt tagtggcgt tctgacccctt aatggagaaa
46681 taagtttgc tatgtgaata tggatataa ataataatgtt aggtaaaaag tgggtgcacag
46741 tcataatgttgc cgtatgtatgc caatcattaa aacaacgttag tagacaatta actgctgg
46801 gaagacatcc acataccatc aggaattttt gctaaatttt catagYaaaaaaaatgtccc
46861 ggcacgggtgg ctcacgcctg taatcccagc actttggaaag gccaagggtgg gtggtcaccc

46921 gaggtcagga gttcaagacc agcctggcca acatggtaa accactgtct ttactaaaaaa
46981 taaaaaaaaaaa aaaaaaaaaaa aaaatttagac ggccgtggtg gtgcgtgcct gtaatcccag
47041 ctactcagga ggctgaggca ggagaatcg ttgaactcg gaggaaaggag ttgcagttag
47101 ccaagatcac accactgtac tccagcctgg cgacagagt gagactctgt ctcacaataa
47161 ataaataaaat aaaataaaatc atctgtgtt aatatggact atttcaataa accaaaaaaac
47221 atgcgtttct agatgttaagg aggaagataa gaaaactgga cccagcaggg gctgctggag
47281 tgggcagggtg ggtgggtgcc taaaccaaag agcccccttc tggcccatgg tggcctcatg
47341 aagtgttcca tttcaagcc gcttcccttcc tggcaagtct gagatgaaaa gggcagcttc
47401 atctttccca tagaaaaggg aatatcacgt cttgtcttta acatagacag gaagcagccc
47461 ctccagtgtat ttccactgtc cccggagcca caggcggctc ccctccacct ctggaaagact
47521 gcccggcccg gccttgcggg ccccgacttc ccctacctgt gccccctgct tgacaggcca
47581 ggctgcctcc attcttccag cggccccat tccgtccca gggccttggc gcccgcctggc
47641 ccctccccc aacctggaaac aggccatccc tatccctaccc cttccacat cttccgttg
47701 tcgtgagcct cWgctcccgat gatatcttca gtgcaccage ccctcgagctg tagaaaagtgg
47761 cctgccttcc cctacccttc ttcccttcc atcttcttca ac tatttttattt ctttcattgc
47821 atttgttattt atttggaaatg atccgtttt ttttactt cagatcgctc tggcctccc
47881 ttcccagaat gcaaggtagc agggttacgt cctgtcatgc tcacagctgc acccccttggc
47941 tgacatataa tagacgcccc ataaacaccg gccccctgtat gaacggcagc caggggcacc
48001 agacagagga ggatggctat gcccagaatc cacccttcatct cagcagctga acacaataa
48061 ccgcacaccg tggcccttca ctacacccat ggcgcggggg gggtcggag gggccggagcc
48121 tatgttgcag ctgttacccat tggacacat ggcttccaag gtttccgtgg caggggagac
48181 agagatggag gggactccaa gctcttacca gccttggcc acaggtgacc caggtcactt
48241 cccttcacac aggccactag gcaaaactgg ccattgcagct ccacaccagc tacaaggag
48301 actggggagg ggagcacagg ggtgatcgat gaggctgaaac tgcctgtgcc accccctggac
48361 cccatatccc agcccaggct ctggccaaaaa gagatccaa aggagctct gcaggcagaa
48421 gtcaccccaa atgatggggg aaactgaggg agtgcaggc tgccagggtg gagctgtgcc
48481 tgcactcagc ctggccaccc ctccacccctt cctcagagcc tgcatccctt gggccccac
48541 catcaggggcc cccaccaggct cctcccttcca gggctcgacg gccttccca gggcacagc
48601 tggccacca ttctctgtcat ccagttcagt gggcagctca ttgcccaggg ctggccccca
48661 gacccggggc tctggatgtg ccgttggccca gccccacccctt ctctacagct ccagacaaac
48721 acctcactgg gagccctggc ccagccggct gctgccacgt ctgtcccttcc atgcgtttgc
48781 tttcgcttca gccacRggga gcccctgtgac cagctggac ttacaagagg agtgtgacaa
48841 gccactggcc aggccttaga gggccacttg gggttgtgt tgccaggcccc catagcaccc
48901 aagttcagag acagggcactg gttaggcagat cctctgtgca cagggccaaac cctgcctgt
48961 tcagcagcgc agctgccacg cacactgcgg cttccatcca ggtggaaatt aatccaaaca
49021 aagtactt cgggggtggct accgtaaaaaa agaaaaaaa gacaagagtt ggtgacaatt
49081 tggagagact agaatccctg tacactgtt gttggggggca aaatggcaca gccactgtgg
49141 gaaacacgctt gggaaatattt aagaaagaat taccctgtgt cccgcaattc
49201 tacttatggg catatagtc a agagaacgg a aacagagtc tggacacaccc acgttacagg
49261 cagcactatt cacagtagcc catagggggg ccacccggat gcccattggat ggctggataa
49321 acaaaaatgtg gccgatccac ccacacacgg gatactggc ctccctgaaa aggaaggaca
49381 ttcttacaca tgctgcaatg tggatggccc cgaggatgtc atgctaagtg aaataagcca
49441 Sacacaaaag gacaattattt gttgatcca cttatgcgag gtacccagag tagtcaaatt
49501 ctcagaaaaca gaaagtggag tgggggtgc caggagtcgg ggggaggggg gggaaagggt
49561 gaaactcaat gggcacaggat tcaatttttac aagatgaaga gWtctgtggc cggtgggtggc
49621 gatggggcc cagccacatg aatgtgttca cggccacccaa tgcacacgtt aaaaatgggt
49681 aagacgggtg cttatgttct gtatatttca cggcaattaa aKtttttaa attaaaatcc
49741 tcttggaa aataagaat taaaatttttatttccatgtt tcaaggacac cagccacact
49801 tcaagtgttc cagggcccg gtggctactg ccccgaaag ggcgagacga gcatctcccc
49861 agcaccgcgg gaagctctgt cgacacagcg ggtcttgaa ggtcaccctt tctcagaagc
49921 tcccacttgc tcaatggccc tccctcccaaa gaccacccaa ggttccgaca tagtttagat
49981 ctgtcctcag gactctccgg gcatgcagag gctctcagtg aggacgcagg gcattcatcg
50041 tcaggcaay gccatgttcc aggcacggcc ggccttatac tccctcgcccc acgctgcgt
50101 ctgcgttgg gagctggggc tcaatgttca ggcagagacc ccccgctgac cttccgcgc
50161 cggctctccca Yggccgggca ctctctgcca cggccggccg cctctgcagg acaccccccac
50221 tggcccatg acaagcacac tggcaacacg tttgtctaat gatctcgatc tcaacttctgt
50281 gcaacagggg acaacagaca cggccaaat actttcttca cttctgagac ttgtctccca
50341 aacgggttc tctgttggta taaaactcat ctcaattttaa gaatttttaa atttctcatg
50401 ctttctttaga aacgaacattt cccatttttcc aaaaatgggg gtacggctgg atcaacttaggg
50461 ttggataaaa gttggayggc aaaaatggggaa gaaatgttca gacatccttcc aaccagctgt
50521 tggacagag agaccctggg caccaggaga aagggttgc tgcctctgccc agcctcatga
50581 aggactcacc acggggccca catttggtca gccccttgg tgggtatca gaagtgaccc
50641 ctccctggact gtcactaaac agtgcactg tcatgataac aaaaccacat tttatgttgg
50701 aactgaaaac gtgttccatcc tactgagaga ctgaaaagca catttactga cattatctca

50761 tctaattcaca tcaaccccca cgatgaaaac actggcccattttacagatg agaaaacYaa
50821 ggctcagagg acRggaggcg cgcccccaac gacgcgcctg gaaaaggcga ggcctggagt
50881 YRccatcgga ctgttaattcc aagcccagtc ctttccccgc tgggtggaa ggcgcataaa
50941 tgtagggga agaactata ttgtgaaggc agccaaaggat tataacctgc cagtctcaca
51001 gacctagcat gaggacttcc caagataatt tagcatggca agtgcgcagc atggggccgg
51061 cacagtcaca tagtaggggc tcgataatgt agccattatg gtttatttgt tattattatt
51121 attattatta ttattatttt tgagacggag tctcgctctg tcactgaggc tagagtgcag
51181 tggtgcgatc tcagctcaact gcaagctccg cctcccaagg tcacgccatt ctccgcctc
51241 agcctcctgaa gtagctggga ctacaggcgc ccggccaccac gcceggctaa tttttgtat
51301 ttttagtaga gacgggggtt caccgtggc tcgatctct gacctcgta tccgcccacc
51361 tcggcctccc aaagtgcctg gattaYaggc gtgagccacc gcaccggcc gttattatga
51421 ttattgcatt aaaggtacgt ttcaaaaata agatgagggt aagtgtgaca acgctctccc
51481 aaataaaaacg aattctggcg atgcagcggc ctteggcccc ctgtacggat aagRctgggc
51541 tccagaggcgc tagctgactc acccaagacc ggtcaagagg aaagtggca tgaagggttgt
51601 gtctctgtcc agtgttcaagc tcactcacat cccctctgc ccctgaggta tcttttcacc
51661 agggccctgg ggactctgc ccatgtgtca gcttggctcc tggaggggct caccctgg
51721 ccggaaaagca gagctggca gcagatcctg agcagaccctg caccacccctgc tgcctgcctc
51781 tagcccgcc ccagaggacc tcattcatcc acagggtgtgg acacgcagca cgggggttcta
51841 gaacttccca cctgccccgc ttccacacca cctccagcccc tccctaagcc cgggcttcct
51901 cccacaaaata ccccccgtcc cctaggcccc cactggcagc atctcacatt ccaggccgtg
51961 ccctggctct ccttctcaacc cccttcttt ccctgcagga ctcagccaag ccagtctgt
52021 cccctgggag cctgcatttt gaagactggc cgcggccat ccccacatg gctctgacca
52081 catgaggcag agagaggaaa ggtctgcaca gcctccaccc acacacgggtt cagcatcctg
52141 gcacccctgt cttaggttgt tgggggtatg gaatgagggg gagggtgttctttaatctct
52201 tagtgttagg cagaccatgg gcagggtggc agaagacgaa ggtggcacaat agacacccct
52261 ttcaatagcc cagactcaca gggagaaaatttcccttcc ccaattccctc ccactgYgag
52321 cttaaggag gggccctggg ttgggtgtctg gcctctgta taatcaYgt gaatctctac
52381 tttaaacaga aaagggggtt ctgtttggag cgagaagaag gcagggttat ccagctgggg
52441 ctgaataacg tcactttatg tcttactatt atattagctt ttcaggtaaa ataagggtgc
52501 ctttttaaat gatTTtagt taaaagtgg caaacatcaa gaaaacatata ttttcaaata
52561 acgggacatt gaatgacatt gtcattcaaa cgactggagt ttgcacaaac ccgccttagc
52621 ctgtgaagag cctcaccaag ttaccatcac ccacgatgtg cagacgagag aactgagggt
52681 agggatggtc aacttgtctc acctggggac ccagcaagcc tccctgagac tgggtcactc
52741 acctccggct aacaaggaaa tcggcagagc ctcaggggac ctggcccagc tgggcccctc
52801 tggcctgccc agccccctgt tgaccctccc cgcggccaccc tgccccctca gccacccccc
52861 tcacaatgtt catgagtgcc agcaggactact tctgtacgtg ctccctcccg tggactgca
52921 ccacagatc atccacgccc agcaggaccc ctgtgtgttgc tcatacgctcY gcagcatgccc
52981 tgcgtgcctt ccgcctcgag ctgggtggcgt gtcctcttgta gaYggccagg gcggcgtga
53041 ggctgtccag gctgtccagg gaggccccctg caaggagagg acaccgtt cagcggcagg
53101 gcaaaccac ccggacaccaa acctaccagg cgcggccatc gccatcctct catttaaac
53161 tcatacgacc ctgaaccgg ggtattatttc cattttacag agggaaagcga ggtgcagagg
53221 aagatggcag agacttgcca agctgtctca gacacagcag ggcggccacca ggcacccctg
53281 tgacttccag gaccgctcg tccccagcccc ctgggtggaa ccccaagccac agcaggggg
53341 acagtggcct ccacaYgatg gatcagaccc gctgagactg tgagagaccc agcccaagcat
53401 caccgtctgc cttatggagc ccagggtctgg cattcagccca tggagaggc gggAACCTCA
53461 gcccgcactg cctgacgtg gagggtggcag agccacccct tgcccccact gcccgtatgt
53521 ggagtgggtt gagccacccctt ttcctccccc aagggtgtcc aggtggccaaa ggactgggt
53581 cccagggttc atcggccatgt gttatgactg gggaaacccca ggcggccatctt gggctgtca
53641 tggccccccat tgcgtcccttgg cggccctccca agaacttttct gctaagtccct ccactgaagc
53701 gaggccccac atctctggc cagggtccccgg gccccctccca caggtaacgg gtaccgcctc
53761 tggcctgatt gaattgtgca gagaccaaga tgaaccccttc tggcagact cccgtcgact
53821 cggccctggg caggcactga cctctggccc ccagggtgtg cccctggca cttggggctg
53881 cMatcctggg gagtctgccc cggatatcg tgcctgccttc tgcgtcgtgc tgcgtctgt
53941 gtgttctgtg tcctgegggc ctggagacca ccctggacca ggcttcagtt cccagccctc
54001 cactgcccgt ctgaaggcca cagcaaggct ggccaggctg acctgcagca tcgatggcag
54061 ccccagacca cactgcccac cggcccccaaa gccccagctg cctgaaagcc cccttccct
54121 tgcaactggc tctgtgtctt agccagggtcc cagccccagg gtctccatgg cattcccc
54181 gcagggtctc agagactctt ggccccggac agacccagggtt ctgcagggtca tgcgtacatc
54241 actgacccac agggcggtgg ccaggccccgg agggccctgc agctggggaa gaccagcg
54301 cctgctgcac cacctctcta caaaaggtcttcc tccagaacaa cactttcac ccccaacgcc
54361 tctctccctgc tgcctcttcc aggccaggac aggcacttcc tccctgtgag ggtggccag
54421 gacaggccct ggaggcgagaa caaggccccgt gcaggcgccca cggccactcg gacagtgg
54481 cagcctggta ccaggccctc atccggcaca acgcttcctc acccaagtgtc gacagcaggc
54541 cactcggtgg acRtggccat ttccgttttca cagatgggaa gaccatggc atccagcat

54601 ggccgactct cggccagaag ctcttcctt gcaccctgca gcctggctca gacaagacag
54661 actgaacatg aagtttcagg cacaaaacctg cagcccggt gtgggcaaga ccaagacaga
54721 cagtggaggc gcgaggcggg gggtagggc tggaaagag aagaggata agtggggcca
54781 gggctgagca gctgctggg ttccagtacc ttccggggc cttgtccac cccagctcc
54841 acagtctgt gtccccctcg cgcacacaca catgcacaca cacacatg cacacatg
54901 tacacacaca caagcacaca catgcacaca catgtacaca cacaggcaca cacacatgca
54961 tgtacacagc aaggacttc cagggccact ccagggacag gggtgccc ggcaggcctg
55021 cgggcaggc tgggggagc cgagcagtcc acttagctgc ctggatgact cagggccatc
55081 catccgttc ccagagggaa cgccagcaac tcaagcctga caggggggtt gaggaatcgc
55141 cactctatTT ttgtacatTT atttcttgc atttcaacg aacagagaca gtcacccctc
55201 agaatcacct tgggaaggag agtcatggca atttccatag gaaaactctt ctcctctgg
55261 gaacaggccg ggcgttaccc tctgcccTT tcaaactagg ggtcacacac gggcggtgcc
55321 ctcggcaaaa ccagaccaca gaggcactcg ttgaggtcac gtgggtccat ctgggttgata
55381 gtatTTgaaa actggggat ttcatgttt aaaaatctag aacatccagc ttctcttgag
55441 gaattaagag atggggccac cccagggtt gcccKggct gtgaatttag gggccacacc
55501 tggagacaaa acacacaggg cttggctcaY caggccacact gctgtccact gtcaatgtca
55561 ttctggcacc agaaggatc acatgttc ctccctggcc taacccaga gtccccatgc
55621 agggaccctt ccagacattc cccgttaggag ggcattgggtt ccgcattcggg acagagctgg
55681 gctggggcac Rttttgggc tgcttctgca ggcaggcag agtttttaac taaaactgtgt
55741 gttcaagcc cattaaaaac agcaaaggTT tctaaagatt tagaatctga atttcaccc
55801 aaatggaaatt attattgttG ccatcattat ttaatctgta ataatttKgg ggtgaaatt
55861 ggaaattatg ttcttagagSa tccctccacc caggaccat gcctagagag acggcccaag
55921 gcccaaaggc aaagtgggtg aggagaggaa gacagggggcc tggggccacct ggcagggtgac
55981 agccacgccc tgcgtctggc ccctggaaac ctcagtgggtg ctcaggctgg acaggacaag
56041 gggagaagcc ctaagagcag gggatctgc aatgcctca gacccggc aacaggagct
56101 ctggcccaagc cccctctgt tcctctctg tcccaacaag accccgcagt caggctqaag
56161 ccccaagctt tcaatcaca gtcacccctc agttttcccc acaaataagat tgaccctgac
56221 cctgtcaatc acgagtggca ctcgtatTT tcccaatTT ggcttaaccc agggcatgaa
56281 gagactctga ccacatgtc gcccTTgcag acctgytccc caaatagagc cttgtatag
56341 ggagctctgg ccagtggcc tggcgtggc tgcgtcgtc ttcctagtga tccctttgaa
56401 gtcaaccctt gccttggcc tgccaccc tgcgttgc ctaccctggc taacagcccc
56461 tggccactt ggtcccttca aggagaaga gggctatTTt ctgccttccc ctcctccct
56521 gggagctccc tggtagccc agccctgtc agctgtgggtg actggagct cacacagct
56581 ctgctgtatc ctcggcact caccggcctt ggtgagtgta tcagcgcacc catctctgc
56641 agaggcagga ggtgacact gcccTTggcc ttcccttccct ctgtgtcag cttcccccac
56701 aagaaggcac tggggccctt ggggtgtc attgaccaag ccccaagat cataccagg
56761 tgggtcagc ctaccctgaa gggagggtgc tggcgtggc tccatcttt taaataggt
56821 acacagact ttcgagttcc tggattttcc tggctcagg gcccTTttt gatggcagct
56881 ggagatctt tggctcacat cgtggccaaa Kaatgacact gtggaaagca ctccagagcc
56941 atccgtttt agagtctta gaatggaaaca aaaagactt atcgaagagt caatcaataa
57001 acagggagag gatggaaatga gggaggcag cagttctgtc cagggattc catcaggga
57061 acgggtgggt cacagtagcc acgggctcat ccacagtgc acagtgcaca gggcacgtgg
57121 gcaagggctg caactgcac catccatggc acaccccgag tctccaggca ggggtgtgt
57181 cacagccctt cacagtggca gtcacccatacc cccagcagcc atccatccc ctcgccaagg
57241 tggaggcgtg tcaccatcag catatgtcc tatttcagca tcaggaaac aaaccaagga
57301 gcaggaaat tacatgactY ggacaagcca caccaccaga gaggaggagt gggcacgcgc
57361 ccaggcctct gaggccaacg ttggccacat Rcccggagcc cttctgtact tcaccagcac
57421 tgggggtgt tggcgRgaa tgagaatttg agaaacaatgatggtagac tagagaaSag
57481 aaaacctggg actgagttcR caacccctggg agcatgtgg gatgggtcgg gctctgaggg
57541 tgggtacagt gggcagggtgt gggatctgtg ggcagggtca gtggcaggc acgggctctg
57601 tgggtcagggtg tgggtctgtg aggagggtgc agtgcaggc cgtggccctt gaaggcaggt
57661 gcagtgcacca ggtgtgggct ctgagggtgg ttgcgttggg cacgtgttgg ccctgaggc
57721 gggcagggtg ggcaggcatg ggctctgtgg caggtgtgg gctctgaagg caggtgcagt
57781 gggcaggcgt gggccctgag ggtgggttca gtgggcacgt gtggccctg agggcgggtg
57841 cagtggcagc gcacgggctt tggggcagg tggggctcc gaaggcaggt gcagtgcacca
57901 ggcgtgggtc ctgaggggcag gtgcagtggg caggcgttgg ccctgagggtt gggttcagtg
57961 ggcagggtgt ggcctgtagg ggggtgcag tggggcagggt tggggccctgaa gggcggtgc
58021 agtgggaggc tggggccctt gggggcgggt caggtggca ggtgtggcc ctgaggggcc
58081 gtgcagtggg cagggtggg ctcgtggccg ggtgcagtgg caggtgtgg gtcaggggca
58141 agtgcagtgg gcaagggtgt gctctgtgg caggcagggtt ctttgcgttgc ctaagggtgg
58201 ctgggtgtt cgttccctt gtgcgtggc ctttgcgttgc cttttctgg gtcggctct
58261 tccacctgcga tggccctttt gtcgtgtccg ctgtgtccg cttttctgg gtcggctct
58321 gacccagaag gaaggaggcc cgtgcctcc catgggactc agacttctt aaagacagga
58381 gcatgttctg ccccatctg ctcctgcac agccacactg cgtgccacta ggggacgaca

58441 gggacagggg ccggggcagg gatgcttaaa acctgaggac cacactgttc agtgacacgg
58501 ccgactgctg tcccctctcc atgggatggc cgccccctca ttgcctcttt ccttcaggtg
58561 tctgctgaaa tgtcagctcc taagggagaa cttecttattt tcatacccaa attgaaatca
58621 gccctccccct gatttctccc gacagcaccc cgtgcetctt cttctgtga cactccactg
58681 caggcaggtc ttgtctcatt tgtgaatttg cttgggcatt aactgacttt ccctctcaga
58741 tagagaaccc tggggatgg agaccatgg gtctgggtg cctcagtgtc cccagtgccc
58801 cacaggtAAC cgcagggtcg tgaggggcct ggagccatgg cagagaggtg ctctggaaac
58861 agagggccca gcagaggcga gctggggccct tctggggagt tggaaactgg gatagaagga
58921 gaggcagggtg cagcaggccc ggagggccca cactgggggt ggggagccga gaacaggagc
58981 aaaagtgtgg tggcagttgg agaagcatga gatggaacaa aatccagcta caaattactt
59041 gcatatggca taaccaagtc caaatgacca ggaggagttt aaacaagtgg ggcaaaagaaaa
59101 cgtaccatgt Mattgccaac caaggcattt aaggagataa actgataaaa caaacaactc
59161 atggaaaaaa atacaaaaga tacgaaaactt tatgctccca aagcttcaac ctggtgcaag
59221 gaaggagaga caaggccaca gtgataggag cactgtaata cattattcc agacaagtct
59281 gatacagcat tagtagggg tagaaatctt ttgaattaat aggaaaacctt aaYgttcaaa
59341 cccaacagcc agaacacgtg accttgtcaa gcacacacgg gagaagttcc acattttacc
59401 acattaggac acaactaaaaa aactcaaaag ccgtgcccac agtcaccatt actgagacac
59461 attaggacac aactcaaaag ccgtgcccac agtcaccatt actgagacac attaggacac
59521 aactcaaaaaa tcgtgctcac agtcaccatt actgagacac attaggacac aactcaaaaaa
59581 tcgtgcccac agtcaccatt actgagacac attaggacac aactcaaaag ccgtgcccac
59641 aatcaccatt actgagacaa aatcagagat aacaataaaaa ccctaaaccc taaacatctg
59701 aaattaaaaag caccRttcta ataggcctag actaaataga aaatatacat tccactcatg
59761 gctcaagtag aaaacctgaa gaaaccaata acctcagaag aaaYtgaag agcagtcaaa
59821 gaaatgcttc catagaagag tttaacccaa Yccacaagga acagatactc ctcatctcac
59881 ataataataat gcaaaagatca agacttcttgc aacatagaa taacccgcRa gaaaggcagag
59941 tggaaaaaaag agcttcttag ctataacagc aaaactccca agtatctaag aacaaacaaaa
60001 aactaatgaa tggtaacacg ctacacaaaaa agttagtataa tatagtctaa gRtctccaaa
60061 aacacaaaaag aagacctcag taaatgaagc tataatctcc atcaaataat tggtaaaaata
60121 gtccctctcc aaatgaatca agaaatttgg cataatctcc atcaaataat tggtaaaaata
60181 attgttaaaa ttattctccg tcaaataatc tccgtcgaat aattgttaaa attattctcY
60241 gtcaaataat ctccgtcga taattgttaa aataattgtt aaaattatttc tccgccaat
60301 aatctccatc aaataattgt taaaacccaa caaagtgtt ccaacattca tccggaagMc
60361 cacaccagac agaaaaataa Ygaaaaagaa tgacaatgag gaggcatatt ttttcaaat
60421 taatgcagtg tattataaaa taatattaat gaaaaYgttag ttgtgtcaat gtcataaag
60481 aaacagatca atggagctga ataaaaacttc caaaaacagg ctgggtgcag tggctcacgc
60541 ctgttaattcc agtactttgg gagggccagg cgggtggatc acctgggggtt gggagtttga
60601 gaccagctg accaatatgg agaaaacccctt tctctactaa aaataaaaaa ttagctggc
60661 atgggtgtc atgcctgtaa tcccagctac tcggggaggtt gaggcaggag aatcgcctga
60721 accSgggagg cggaggctgt ggttaagctga gatcacgcca ttgcaactcca gcttgagcaa
60781 caagagtgaa actccgtctc aaaaacacaa aacaaaacaa aaaaacttcc aaaaacagat
60841 ccaagaatat gtggggtttc caggttaaaa tggcRgattt aacacagata tctgattttgg
60901 ctcccttctca gaatcccatt aaaagttcag taaagaaatt taaaaaaaaat taaaagtcac
60961 aagcctcYaa ggataggSag agcaggaggg gaagaaaYaa caagggact gtgggaggt
61021 gggaaacagat gggcaagtgg taacagactt ggtacatcaa gaaagctgaa tcccaggctg
61081 gccatgaaga cagccacatt ctacaaagtc tcccggagaa taagaagggg gaaaatcagg
61141 ttacaggaga ggtcatagac aaagaattgg aatcagaatg actcagagtt ctcaaaagca
61201 gcattcgtatc caggaaaaca gtggaggaaa gccttcagaa ttctgaagga aatgattgt
61261 ccaactggaa ttctgcattc agacacattt cccatcaatg atgacaaatgg aataaggaca
61321 ttcttagaca tgcgtatctt caaaagattt ggcaacaaag aaagagtaag acacagat
61381 tggaaagggtgt gcaccatcaa aaacaaaaaa ggcaacaaag aagagtaag acacagat
61441 agaaaaacgga agccccacac agatagagaa aacggggagcc caacacagaa gacatggct
61501 gggaggccca tggagggtgg cgagaagtga ccctgagggtt atggctctgc accagaYgtg
61561 gagagcgaact gagccatact gagcagcgtg acccacgggg tgaactgtct tcaccaagac
61621 cccacctcca tcaggccatc tgcgtatctt ggacaaaatg cccaggagag cctggcttag
61681 gttcttggca gtacttggct tggccatgtg cgcgtatctt cctctgtca
61741 ggaacgcacaa cattcacaga agttccagc tttcccccac ccgagagctg gaggcaggcc
61801 gtggcgagct tagaaggcagg acgtacaggt cagccccctc cctctgaccg tatttccgt
61861 gcagggtccag cgcgtatctt ccacccggc cttaccagggt gcctgactgt gctgagccca
61921 agatccctc gagacccatgc tcacacttc cccaggaggg gcctgtcag ttcccccggaga
61981 agcagaagtc agaccacgc ctggagactg atctgctcac ctccatccag
62041 cggggcgt gccgcctt ctttacactg ccagcgtgt ccagctctcc Rtttaat
62101 taatcgtctt tgatattcat ttgcattaa ttataaaatag ttgcataatg agcaatacag
62161 tgaataaaaat tagtaaatga tacaggtagc aaacgggtct agggaaacggc ggagggtggag
62221 caggagggga caacctggca gtctggcact actgagggga tcatgaggaa gaggcaccatg

62281 cggtctat cctaaagata caggatgtct atcgatagc ccagtgcct cttggactg
62341 acccttggcc catggccat tggtcgccc catggactgg ctgcagatg gcaaaacttg
62401 cacaggctg taaggaggca atggggact gccccttgc gcattgttgg gttggagtga
62461 gggagctgg aaacaacacg ggggtcatca gtgggtgaga ggagaagcaa atgtgggtgg
62521 ggcatcccac aaagtgcac caaggttaga agtttgcata ctacatagat ggatggatct
62581 taaaaacaga gtgccaagtg caaaaggtaa gaaacagaat gagacacaat tcagagcgcac
62641 ttacRtaaat tatggataca cagacgcac acagcaacac tcatgttgc aaaaatgtaca
62701 ctaaggaaag qatacggtt qatcacctga aatggttgtc cactgcaggaaatcaga
62761 aggaagagag caggctgggg agataaaatgg cttttgttc ttctttgtt ttaatgtat
62821 taggacaact gtttagaggg agggaaaggt caaatttga cctcatgcct taagcaaaaa
62881 cagattccaa gtggatgaga tgcataaaag gaagtcatgg aagtgttgc agaaaaatgt
62941 gckaggcacc ttttatcat cttgtgttgc ggcgggctt ggttccaag cctagggctc
63001 tttcccgcc ctctccctag ctgcggccatc ccctgtgagc agtttccac tctgtggac
63061 tctgtgtgtc gaatgcacgg gcacatggca cacttgcgt gggatgggt gggggggat
63121 gggcacatc Wctaggggat ggcgMaaggg gggatgggc acactcaacta gggatggcc
63181 tgaggggcaga caggcacact cactgttag gggatggct gaggcagac ggcacactca
63241 ctatgttaggg gatggctga gggcagacgg cacaactcaacta gggcactg cacaagaca
63301 gaggtgtgt gctactgcct caacctgcctt accatgttgc tgggaaacaag tcatccccc
63361 tccatgtgtt aggagtcccc aaaacgttaa ggcacatgcctt cgcggcYgcag agccagaact
63421 gccacaStgc ccagaagagg agagggaggc tcaggaaattt agactaaggc catgtgttgc
63481 gcaaataggcc aggccagaat ggcagccag ttcctgcaca caagactgc cagggctgc
63541 agggggccacc accRggcata ggaggacca accgcctggc accagcagcc ttcctccaa
63601 catgaggagc cccacccagg ggaggtcaac gggacaagg tcatccccctc catttcttac
63661 Yaaccaagag agagtcgtc tgggtgttgc tattccctta aagataatca ttaagatgaa
63721 atatgtttat gtggaaatgtt aatcaggaca cagcagaaag ccctgtgttgc aagcaggcag
63781 gcttttggg ggcgtctggg agcagaaggc aggttcagcc ccacggggc catcagcaca
63841 gagcccccact ggtgtacttt attaaaggat gtcatttgcatttccatc cccagccctt
63901 tataacttta ttactgtactt tgaattaattt tttgcactt caacagacac agtcatcca
63961 atgcattttt tatgacctga ggtggatcg ggggtgggaa ggggtcctg gggccggct
64021 ggcattcagg agcgcattcga gaaaggctgtt gccgaatctc agggcagtc ctgacagttc
64081 tctaaactcc caYgggatgg gctgccccgc ggacagatgg acgcctccgc ccaggctcgc
64141 tcagcctcag ccRtgcattca gggagaaaaa ggaagcacta aaggagagaa ggcctctcgc
64201 ggcagtctt cagaacatac cccgcacag cagaggacaa aggacacccc tcagtggttgc
64261 acacagaagg tccctggccctt cccagcacctt gcacaccctc acagagccac ctgggggttgc
64321 acagcacggc ggcctgcctt gtcgtggccat tacacccccc ttgtctgttgc accaccctcc
64381 cctatttgc cccatattgc ctcatgggSa gccaatccca ctgcacccca gotccaggag
64441 tgcccaagg gtcctgttag gccactggaa cctgtatgtc tgggtgggt tgcacaggaa
64501 gactccctgtt ctttcccttc tgcgtggaaatgtactt gggRtgcata agtacccct
64561 tcccccaagg gcaaggcaggc gcaaggatgaa gggcagcccc ttgtgtgttgc gacactgttgc
64621 aggtccccaa atccagctat ggctgaaggat gatgcacccca ttgcacatccctt aattcttcag
64681 gtcaattcat ttagctctt gtccttgc accacgagag tccctgtgtc ctttacaaac
64741 acagacacag ccaaattgtgc tggctgaga gagacaccat ggttgggtt tgcagccaca
64801 gctcttttgc acgatttgggg catcagacca ggcactctc ttctgcctt tgcctgcacc
64861 tccccatgtt gggcgaggag aaggctgtca gtattttaagg tctatgacca gaaaaggggc
64921 caggctctgg tccggctccc agatacgtt cttggggaa tgggtcccag cagaggtgt
64981 aggtctgtt ggttggggcc aaggccaggc ggcctatgtc agtgcaccgc agctgtgcga
65041 gcccctcttgc ggtgtccctgc agcaggctg gggcttccctg tctacagccc tcacagaggt
65101 tcccaaacac aacttcccccc ttttccatccc agggcttgc cctctgttgc acagccccc
65161 cagagggttcc caaacacaac ttcccccttcg tccccccaggc accccccgaa taccgcaggc
65221 tcctctgcac ctggccctgtt ccactcttca gggccYgtc cagccactc accttcaagg
65281 attcaggggg gtctcggttgc ttccatggc tggggccgggt gacaagagcc agcaaggct
65341 gcccctgttgc agcatgttgc taagccaggg agggccatcca ggactcacca cagccccc
65401 ggctccaaggc ttccctcttgc ctggcagtt gggccaggaa agggagctg gcctgggtaa
65461 tgtcacgggg gcacccgcagc acgcggcagc caagtgggtt gggccatgg agaggccggc
65521 aggttttttta ttgaaaat tcaatgttgc ttccatggc aaaaacggat cccacgggt
65581 cccaaaggaa gggagaattt ccagggttgc tgggttgcac gcagaRgagg tggggggccgg
65641 caagtgggtt gctgacagcc agtgcacccgg gacccatgtg gcccggggct tggccagtt
65701 ctcagcagat gacaggcggg aaatccctgtt caccatgcYgtc gtacatgtc cagagccct
65761 gatttggaca tcatgttgc ggcgttgcattt gtcctccggg aggcctcagg tacctcac
65821 tcagaagggtc caggccctgc tgcccttgc tggccggcgtt gacccatccgg ctgcacttgc
65881 agctgtatcg tccacccatgg ggttggggcc cgcgtcggc caggccatgc ccaggagcag
65941 aacctgcctc tccccctccc tcaaggcag ctccctctc actccctcagg gtctctggaa
66001 cctacccctc cagcgttcat gcccctgttgc tggccacgc tcctgtacttcc agcagcatct
66061 ggcctcccc catctctgc ttcattcaat ccaaggatccca aggcatccat ttcacccat

66121 accttcagggtt cccttgcgtt ggccttcacg tcctttgatg ttgagccctg cgttcatgca
66181 ggggtggcggt gggccctgca ggggtgaggt cctggcccccc tgggaagagc ttgcagccaa
66241 cacaagcggg aggtctcaga tgctgccccca ccctgtgcaa atgaggcccct ctcctaccgg
66301 gccatgagcc acgtgcaggg actcagcacc tgcctctgtt ccctccctcttcc tccctgacccc
66361 tcttaggtcag gagtccaggc acacagcagt ytggggtgc tgcagccaag cacaggctct
66421 tcctagctga ttgtgggacc tggggcaggc ctgcacgggt gcccacactg tgccaccaag
66481 ctgccttcctt aggagaaga gttccctcaga tgcccaaggca ccaggacttccacacccctc
66541 ctgttttttc ttcatgttag ccgggtggat gggcacgggaa cctctttaac gcaaattgagg
66601 gaacaagecc tggggccagg caacaggcact ccctgtgttcc ccccaagcact tggtctgtc
66661 ccccgatcc cgatctcca atggcccatg gtgggtatttgccttgc acgagccaaR
66721 agggagctga aagtaccagg gccttggccc agctgaggac cttgtgtccca tcacccacca
66781 tgacccttgg gcacaaagac ccccaactt cagtcgtggc accaagcaca gtccatgtct
66841 gaccaggacc tgctcaagga gcaggagtgaa cggaggacaa cccggggac caccggc
66901 ctcccccattt ctgggggtgg aagggtggg aggcttggg cctttgtatg ctcccatgcc
66961 ctgatgtccc atccctcagg acgctcccat gcccctgcgc ctccctgagg ggcaggaaaa
67021 ccccatgtgg gccccgagca gccaggctcc ctggatctgc ccagagaaga gccccccgca
67081 gaggtgtgc aatgtctcacc acgccccctt cagctcaaga cccttggaaact tgccagaccc
67141 tcagccctctt ctcctcaggta gtgtcaggga ctccttcc cccctcaggct gctggggccc
67201 caaacacat gtcctccccc cctctgttcc agccccccat ggggtgtagg agcaggctcg
67261 cccaaacctt tttcttggtc ttcctgttcc ttcaacgctg gctgtgggat ctctcttcag
67321 acctgcctt cccacgggtt aatcccaagg tgcaactca gagacagcac cgagaggccc
67381 ccagtgcaggc cgcactctg agccaggatc ccaggccatc aaacctgtca gggacggggg
67441 ttccctcccat ttgttcttcc gtgtgaactt cgtgggttca gaagcctgag tttgggggt
67501 gtgacaattt gcccctggg accccacaca ctcctgttcc atgatgttcc ctgcctttaga
67561 atatccacgc tgcaatttggag gggggaaagg gtccttcata ctcccccacac agcaccaga
67621 ccacgggtt gacgaggctt aatttgcata gcccattttc gagaaggag gaaatcaga
67681 atcagagtga gtgtcgatc ctgaaaattt aatccccaaag gcagcctcac agggccttcc
67741 ctccccaacc ggggcagggt gcaaggatgc actacaagcc cagatccca ggcacccctg
67801 agaaccagga gggggagggg ggcacccctt ctcgcagaga gatggcaaa ggcctgtt
67861 gccggggcccc agctgccttcc aacactgcaaa gggctcagaa gactggacca gcagctggcc
67921 gggcagaaac tatagaaatc atcctaattt tactttaaa gctgggggtt tccttttctt
67981 gcttaaaagg tgggtgtgacc ctcccaggta gcagccagcg gcaggctgt ggctactga
68041 gatggggctgg gtgggtggcc ctctggctg tgctcccaac cagtcggcca tggcttctg
68101 cctggccctgc agaggtgacg tggtgtccat cacagatttgcatgtgtga actgaactt
68161 gccaaggccct gagagggtt ctttttttttccatggccatcggcccttgc gtcgtcccg
68221 aacctccacgc atgttctcc ccactttctg ccaggcaacc ctcctcccttgc ctccttctt
68281 ttgcaacagt tccttacccgtc ctggcccttgg cccatgcctt ccaccaggac aagtactgt
68341 ctcccttcttgg ggttttttttccatggccatcggcccttgc ctccttctt
68401 tcaagtattt tgcttaactac catttgccag ctttgcgtgg caccacctt aggagagcgc
68461 tctgggtgtt caccatcaca ctttttttttccatggccatcggcccttgc ctccttctt
68521 cacacacatg cacgcgtca ctttttttttccatggccatcggcccttgc ctccttctt
68581 cacacacatg cgcactcaca ctttttttttccatggccatcggcccttgc ctccttctt
68641 cacacacatg cgcactcaca ctttttttttccatggccatcggcccttgc ctccttctt
68701 cgaatgtcaactca ctttttttttccatggccatcggcccttgc ctccttctt
68761 gtcacatata ctttttttttccatggccatcggcccttgc ctccttctt
68821 cactcacaca ctttttttttccatggccatcggcccttgc ctccttctt
68881 tgcacttcaca ctttttttttccatggccatcggcccttgc ctccttctt
68941 ctcacgcgtt atggactctg ctttttttttccatggccatcggcccttgc ctccttctt
69001 cactcacaga ctttttttttccatggccatcggcccttgc ctccttctt
69061 ctttttttttccatggccatcggcccttgc ctccttctt
69121 gtttttttttccatggccatcggcccttgc ctccttctt
69181 gtttttttttccatggccatcggcccttgc ctccttctt
69241 gtttttttttccatggccatcggcccttgc ctccttctt
69301 gtttttttttccatggccatcggcccttgc ctccttctt
69361 gtttttttttccatggccatcggcccttgc ctccttctt
69421 gtttttttttccatggccatcggcccttgc ctccttctt
69481 gtttttttttccatggccatcggcccttgc ctccttctt
69541 gtttttttttccatggccatcggcccttgc ctccttctt
69601 gtttttttttccatggccatcggcccttgc ctccttctt
69661 gtttttttttccatggccatcggcccttgc ctccttctt
69721 gtttttttttccatggccatcggcccttgc ctccttctt
69781 gtttttttttccatggccatcggcccttgc ctccttctt
69841 gtttttttttccatggccatcggcccttgc ctccttctt
69901 gtttttttttccatggccatcggcccttgc ctccttctt

69961 ggcacactgc ctccagccca aggtggccag tggtgactag gaggggtcag gtccatgtt
70021 ctgggtctta tttttgttaag gaagttagaa atctggatt gcatgtgaag tctctgatct
70081 ctaataactg gcaaaaaagca attaagaat aatgttaggtg ctgtgtgtc tgtgaaatga
70141 accaaagctg ggtccccagg tctgtctgcc acctcagotc caceccaga ccaaagctgt
70201 gtcctcaggc ctgtctgcc cctccggctcc acggccctgag caaccgttcc cctgtatcc
70261 cctccgtctg gcctggagag gttttgagca cagtaggttc tggccaacca gagtgagttc
70321 acctgcctct gcctcacagt ggctgaagct tcccggccg cagttcacgc cccacttgg
70381 tctccacate cctttccac cttgaaggac agtctctgtat ggatgtgtt tatgttggaa
70441 tcctcacgca ggggcctgag gaacactccc tggttatcca cacgcaact aacactcagc
70501 acgcaccagg ctccctccctc ctctcatgga ctctcaattc agccccRcac agccaagggc
70561 tcacccacat catgtgcgtc tcagacaaag gactaatttc ctaatacat aaagcgctct
70621 tacaatcaa caRgaaacaa tccagcaatc caaaaggaa atggacaaac ggtattgata
70681 caaaagtgtt gggtagagaa gggcatgtc cctttaaatg atatgaaagc agggaaagtgc
70741 atggccctg gctaagggtt ccacccctac ggacctatgt gaggacaggc actcttgct
70801 ttgtggccaa atgtcacatt tcccaagacc acccttggcc accatgtccc catcctgtgc
70861 ctataaaaaac ccgagacccct aacaggcaga cacaggccgc tggacgtcga gaggagcga
70921 tgaggggagg aacacacagg cggtgtggac tcgagaggaa cgaccaacg agcaccagca
70981 cgctgcaggc caccgaccgg cagaagcaga acgacacggg gtttggccgg ggcagtcaga
71041 ggagagccca ggctgcttag cggcccgact ccagggaaag cccttgcac cccatcctcc
71101 ttctggcttc ccccatctgc tgagagctac ttccactcaa taaaacctcg cactcattct
71161 ccaagcccac gtgtgatctg attttccag gacatcaagg aagaacccga gatacagaaa
71221 gccctctgtc ttgcgacaa ggtagagggt ctaactgagc tggttaacac aagccgccta
71281 cagacagcta aagtaaaaga gcacatgcagc cacacgccc ctggggcttc aggggctgca
71341 aacattcacc cctagacact gctgtgggtt cggagccccc cagctgccc gtctgtattc
71401 tccccccagag gtttgagcag aggacactg aagaaggaag ccacaccccc atagcacgccc
71461 ctggggagg gacaagagaa cttctctcat ttcaatatga ggggacaatt gtcagaaaag
71521 aaaagacagc acgtttttttt ttctttttt tctctttttt cttctttttt ttttttttt
71581 ttcccttcctt cttttttttt ttcttttttt tttttttttt ttttttttt
71641 gagacagagt ctctgttacc caggctggag tacagtggcg tgatctcagc tcactgcaac
71701 caccgcctcc cggattcaag cagttttctt qcctcagcc tccaagcage tgggattaca
71761 ggtgcctgct atcatactg gttaattttt tttttttttt gtatKtttag tagagacagg
71821 atttcaccat gttagccagg ctggttcaa actctgacc tcagtgatc catccaccct
71881 ggcccccac agtgttggta ttacgggtat gagccaccga gcagttttc agtgtatgag
71941 aagacgctcg acctcaccc taattaaaga aacaacctgc atcagattgt ccaaataaaa
72001 atgtacggat caggggtaca tgcgtgagaga aacggatcca gccaaccatc ctagggttgt
72061 ggtcattctt atcaaagtgtt aaaatgcaca aaccttttga atccacaattt ctatgttata
72121 aatttaatcc tacaatgtc ctatatacaag tagccaagaa aataagtgtt aggatata
72181 Sggctgcatt ataagaaacR gtggaaaagaa gtaacgtccc catcgtcaa ggtctggta
72241 agcaagtgtat tcacaggatc gaatgccttgc cagggaggaa tggcaggc tggagggggt
72301 gctgtggacg acgttccaaatc catgttccct gataatcgca aatgcagggt agcacacgg
72361 gtagtctaca gtgtgcaggaa aaaagtccaga aaaacgaaaa cacagtatctt acaggacaca
72421 aaatccctt agccagagat tctaattccat ggatttatat atgtgtgtt acaNatgtgt
72481 atatatatat acttgtgttgc acaggaca ttctgaaagg acacagaaga aaggtgttgc
72541 gtggctgccc tgcagaaagg agctgatcc caggacagcc aatggggagg ggatgaggac
72601 agcaatgagc atttcatatg gtcataatcc ttgttccagc caaaaagcaa gcccacgcaca
72661 tgcaggagag gccaggactc tccctggc tccctggc tccctggc gtcaggggca gagggagcc
72721 ccaggctggc ctgtctctgtc acgacttac acttggcagc tgaaggggc caccctgaca
72781 gtcagggccc agctgtgtc tgggggttgc ctggcttgc accgatctgc ttggcttgc
72841 tctgggtgtt ggcacggcag tattttttttt gctgcaggc cccctgtgg gggcccttga
72901 gtgtatgttgc gctgtctggg ctggggaccac agatgtgatc ggaccagcat cccctcaacc
72961 ccagatgccc tcYaggactg caggccccag acacagagag ggaacctgac gctcatgggg
73021 agggctgaga ggcaggcaga gggggaaagg ttttctgtac cacggccac aggaccagca
73081 gtgcagaggg cctggggatg agggaaagccc ctgctgttgc aaggaggaggg ggcaggaga
73141 ggagagatgg cccggggccgg gaggaggcaga aagtgccaa gagaacttagag gggggagg
73201 agccctggct tggggctggc gaggaggcaga ggaaggaggg tggaggcag gggaggcagc
73261 gtgtggacaa agactgagag gggggatc acgacttgggt cagaaagtgg ggcagggg
73321 gattccgggt ttctgaccctg ggcacccggc caccctggaa agccctccctt gcaaggagg
73381 ctcggggagga gccgcgtggg aggaggagaa ggcgcgtggg ggcgcgtgg gctgagatt
73441 ctgtggggac acaaggacag acacccatc gacactcagg caggatgttgc tggaggcaca
73501 gagagagggt gggccggagg cccctggat gggccggcc tggaggat cagctgcaga
73561 gcacccgcgtt ggcgcacaca gatcgttgc ctgtatgtaa aagtaggcatt aacgatgt
73621 gccacagccc acaagtaaga cacatagaaa acctgaggcc cacagagggtt acaggaaacgg
73681 gccagggttca cggacaggcc ttttctgttgc tcaatcttccat ccaccagaac ccaggcttgg
73741 caccaggctt ttttctgttgc ctctccatcc acgaggccag aattgggggtt ggcacttgg

73801 cagcctgaag acctccgggg ccctggctcc tcccgtctcc acctgccata gccccaaaacc
73861 agtgtgggtg tgaagcgaa ggagactcYg cagcaaggcc cagctcttat gcccggccact
73921 gccttgctc ctgcctgtat cagtcctctg gcagcaaggc aggtgcctg agacctggag
73981 gtggcccccag ccccagaacc caccatgcgc ggctgggccc actgtgcctg cccttaaggc
74041 aggcaaca ctaaccaaag ccagcggaca tccactgtc gggagcatc tggggaaaggc
74101 tcctgtgggg tcccagaccc ggggcgggtgc cttcaccacg gcctcctcat ccagcatcca
74161 caactcttga gcttaggaagc ccgggtgtcc ttcccagagg aggagagccga agctctgcca
74221 gtctgcactg aagtctgtgg tgcgtgtca cctgagaaca gtcagtcag tctggctgga
74281 tcctgagctc ctccccccgc ctgtgttgc gaggtgacag ccagcacgtg agcaagtggc
74341 ctgggttctg gtcccagcaa agctacctcc agctgtgcac ctgcgtcagg cacctccagc
74401 tgtgcaccc tcgtcaggcc ccctctctcc gtcagacccct cagtcctg atcagtgtcc
74461 acaagtaaga cacatagaaa acctgaggcc cagagaagcc aaggaagtcc ttacaggctt
74521 ctgtcagaga atcggagaa aggagggac gccagagaat ctgggcttaa gtgacttggc
74581 cgtgatcaa acacaggcag agtcacaaac acgagacccc gctgcccact cccgacggag
74641 aggttaattcc acaaacacga gaccccgctg cccactccac aaacatgaga ccccgctgccc
74701 cactcccgac ggagaggtaa ttccacaac acgagacccc gctgcccact ccacaaaacac
74761 gagaccccgcc tgccccactcc gcacggagag gtaattccac aaacacgaga ccctgctgcc
74821 cactccacaa acatgagacc ccgtctccca ttcccgacgg agaggtaatt ccaggtcagt
74881 ggcacatccgc ctccccccgc cccaggccc ccagctgcca ccaagaaaca ctctttgttag
74941 gggtctgtcc ccactgggt gcagaggtac atcccccaag gcccagcccc gcaggaagcc
75001 ctccctctgc agccatttctt gaaatcacct ccaagaaaaaa agaaatgctt cagttccaa
75061 agatataccc aggccaggaaa tggccgcag gcagcatgaa tggtaatta cgtaatttgc
75121 ccaagtccag cgcacatcat ccctcctaag tacctttagg gaaatcgtaa ttaggcgtga
75181 aatggctcat tttcacaact ccctttctc ggggttctcc agggcgcatg cccctgtcag
75241 ctggccggac aggccctctc tgcacaaggct taccctcctt gtttcaatt taattttgt
75301 acctttctc tcgacatccct gaccccgatc atccctgcgc ttggcacac tgcacccaaag
75361 caccatgaa ttactgagct ataagccctc catggaaatc ttttttgcatttttagaa
75421 aacactataa tcaaatccct gagagtccgg agggatgggc ttggcctgccc agaggaaaca
75481 tatcaccggaa ttaagtgtatt tagcagaggt ggtcccgccg ggaggaggcc agggcaggt
75541 gggcagggtc gggcgtgagg tggcaggca ggttgccttc gggggcaagg gggggagggg
75601 tgcgcagagg cagaggatgg gcaatgtggg attaccagaa agtgcggctc gacatccca
75661 gggcctgcag gcttccttag tgcttacccgc gcccctggca ctcacaggcc attgtgcgg
75721 agatgctcaR ctgtgcagag ggtgtcagg tctgggtgtc acttatgcgc tccaggcc
75781 gcttctgtc atgcccaggc aagaccacgc ctcaaaacaca catcaaatac atacgcttga
75841 cgttatgccc gggagggtc tgacacacagg cacacagaga cggtgagact tggggatact
75901 tgggttcagg tactgtcacc tgggtctca ctctctggag cgtgattttgc tctggcacca
75961 ctgcactgcR gtgtgtctgc gtctatacag acttcttccca ttacacgcgt gtacctaataa
76021 cataccacca ttcatgtact gtcaggact ctttcatttt gtaaaacaat agggttatca
76081 aggccctgcgc agtggctcat gcttctaata ccaRactttt gggaggccga ggcaggcga
76141 ttgcctgagg tcaggagttc gagaccaggc tggccaacat ggtgaaaccc catttctact
76201 aaaaatacaa aaatttagccg ggcattgggg tgggcgcctg taattgcgc tactcaggag
76261 gctgaggcag gagaatggct tgaacccggg aggccggaggt tgcaigtgcgc cgagatgt
76321 ccactgcact ccagcctggg cRacagagcg aaactcagtc tcaaaaaaaaaaa aaagccgggg
76381 ggcttaccaa aagatgagag ggacccaggc ctgtattgca catccacate agtgcacgt
76441 taatggtaYg cagacccctt ccccRcatgc acacgcattc tgaataaagc caggcccaga
76501 aggacaaata cggcactctg ccactcacat gaggatctgt aatcaaacgc atgaaagcag
76561 atgacaccat agaggctgcc agagctgggg gaggcggggaaa tggggagttg ctgttcaatg
76621 ggtttaaagt tttagtgcgt ctggacgagt aagtacttga ggtctgtcgt acaacgttgc
76681 gcttgcgtact gaccatgtgg cactgtgcac ttgcggaaatc tctggagaggt ctgagctcac
76741 gtttaagtgtt ctttccacaa aagcaaaacaa caacagcaac agaaatgaag Kgacagaagg
76801 aaaccgtggg agaagacccag tgcgttaccgc ctcgatgtgc tgatggcgcc cccgtgtctg
76861 cacaatcca cacttgcgtga actgtaaagg gccgggtgacc tctacagtcg cttaccatgg
76921 cgagcataaa ctgcagaga agaggatgag aattctgata ataggacaaa tcaaagctt
76981 acaagtaatt aaaggacaaa ttgtgtcata taacaagtc ttccctgtct cgcagctgac
77041 atggatcaat gtccaaaatt ttatccaaa ttgttataaa tacacatgag agtcttctcg
77101 gtgcagtgaa tagtgaagtc agcagaggca gtggcttacc cctgcagtct cagcacttta
77161 ggaggccgag ggggttaggt cacttgcgtt ccagagttag aggctgcagt gagccatgt
77221 tgcactgcR cactcccgatc gggccacag agcggacccc tgcgttcttagg aaaaaaaaaa
77281 agaaaaaaa aaaaagaaaaaaa aaaaagagg aaattagcca atatcccaaga gagtgaaggc
77341 agaaccggta tcagttttg cccaaatgtaa gaactccaca gttgcctccg cattctgtct
77401 tttaaaaagt gtcaattttt cattatgaaa tagagtgcattt cttttaaag taaaatcgat
77461 gctccaggaa gatattccac gtttgcgttgc gactttgggt ccctctcgac acttgggtt
77521 gattaaatta atccaaactc cagtttgcata agccagtgca cctgtgtgcg gctgttacat
77581 tgaccaatg ttttatcaaa tctccgacta tgcccaagtg cggcccatg gaagccacac

77641 gggatccctg acagacagaa acatccatac tgctcttaag gtggctcagg cactgttcca
77701 agcacctcg tctatagtgc ccactgattc ttgcagcag ctgtgtgagt tagctgttat
77761 gcttatcttc attctgcaga tggggaaact gagaaaggtc atcccaattt cacactgctc
77821 ggaagtgtca gaaccagat atgaaacctg gcatggctac caagagtctg tgctcttgaa
77881 cgctgtgta ggcactcct tctggggctg ttacattttta actctgacat gtcaagcaaa
77941 tcactataag ttatttgagt catttcttag ctgcttcattc tgtaaagtgc ggtgacaatg
78001 gctgcagggtc ctgccactca gacgattccc acgcgtgtc gYatcaggat gaggatgagt
78061 agatcggtaa agtgccttgt agactgttat gatcatattt ttgtcttga ttccactgcc
78121 agcacaatac cttttttt ttaataaagg gctgttgaac tgcaattccc atgtttgtat
78181 acattcttct tctgtttctt tggtaaagt ttcaagatc tcttagggtt caaatgttta
78241 aatttcaaattt ccttggctc atctgactaa ttacaagat aaaatatggc atttccaaca
78301 agtctggaga atggaagcac atgaaattat gcacgaccat gagcatctgt gtcgggaggc
78361 gggaggcgt tcaccctca aactcaagtg aaaaagggtt gccgtagttt aaggctgctg
78421 acaaataact actgttctga gaaatcacca ctgtttaactg ggagaatcgt ttttcagcaa
78481 atatatcaag aaaagtaat atcaagtata atttgagat ttcagatggaa ctccacattt
78541 tggaaaaaa aacttctgtc tctgtgtaca agtgcgtaaa acacagcaga gacattttc
78601 ttccaatcc tgcctaaata ctgtacaaat atcatcgat ccatcttcca tgactgtgac
78661 agtgtatctt gtgtgtatgt gtgttatcat attatctgt ttgttaatg ctatgtttt
78721 taaaatgtgt cacattaaaa aacaaaaaca gccccggcat ggtgctcac gcctgtatc
78781 ccagcacttt gaaaggctga ggggggggaa gggtgatca cgaggtcagg agttcaagac
78841 cagcctggcc aacatggcaa aaccctgtct cctctaaaaat atacaaaaat tagccaggcg
78901 tggcaggc caccgtat cccagctact cgggaggctg gaggcaggag aatcgcttga
78961 acccggaggg agaggttgcg gtgagaggag attgtaccat tgcaactccag cctgagtgtat
79021 agagtgagac tccgttcaa aacaaaacaa aatgaacaaa caaacaaaac aaacaaaac
79081 aaaaaaacag aaacacacac tacactgact acactgaacg cctcaagccg ctgagcagcc
79141 atctccaggc ccccaaggac caggcacca ggtctggggg tggccagca tccatttcc
79201 ccttctcg agcagcagcc cagtttattt tccttgcattt ccttgcacca tgcattgcaca
79261 cagccctccc caagccaaac atctctccc tcacagatg tggatttga gccgcaggag
79321 ccagaagagg aaaaggtagt ggtgtacta aacacctgtt cacaatgtt aagaccctcc
79381 tggcaggc ccccgactg ccaggccctt ccacagcccg actccccgtg gactctgcag
79441 gtttcttaat tccctccctt taactcttcc tggctcaggg cagctggact tcatgcaagc
79501 tctctgccc aagagaccct ctctgcaccc aggcgcctc cccatcagat gccagagata
79561 tcaccagccc tcgaggctt acccctgtct tggggggat tcctgtctgct ctgcctgaat
79621 tcaccagggt gggcagtcct cagtgactt ctccagctt gcatggagaa ggcctcacca
79681 tgcatttcag agagaagtgg aggctgcacc cccacaggaa gtggggccca gtccaccgtg
79741 ccctggacgc tgctcatcca ctaggactgg gttccctgt gcagggttcc accctcgctc
79801 tcctctctgc ctgtccaca cctgtgtctt tgacccatc gaccatggac ctgtgtacag
79861 gtggccccc acccacccctc agctgcacc tcccccaggct ccagaccccg cccctgtctg
79921 cctgcactt cccctgggtt gtcaaaaggc accttcatttgc caacaagtcc caaacaaaacc
79981 tggcacctt ccctacccca accatcagc cattgtctgt gtgcctatctt catctccaca
80041 tccaaagggtt caccgcattcc tggcacctt cccatcttcc cgtccatgtat gagctcaggcc
80101 cctccctcta cctggatgca gcagccctt cccagacccg gccccatccg ctgctgttcc
80161 ctggcccttca ctccctgttag ctgcagtcct cacctgagct tgccccattt aaaaatctt
80221 gtacccctt cttccaaaggc ctgcagacc cgtggggag tgcagcaca ctggcttaggg
80281 cctccctgtat ggcctgaccc gcctgccaac cctccctggg acctgtctat gttcccgca
80341 cgccttataa gccatgggtc cgagaaaaca ctgtcaactc tgcccaaat gcctgcccctg
80401 cccttctgg cagactccata ttcaatcttc aaaacccacgc tcccaactcg ttaggcccgc
80461 aaagagagca agttacttctc tctcttagac tcaagccccccctgttccat tttttaaac
80521 cactgtatcat gcaaaaggat aagtcacca gtttccaaat cccctaaacca gaccatgagc
80581 tccagggttta aatcatgca tcagattcat ctctgggttca cctgccttgg ctgagaggag
80641 gctcatgaag gtttggcaca atggatgaa ggaatagaca aggagcaaga atacgagctc
80701 atccatcatgc ctcagaatag catgaaatgc tcagggtgt tgggacagaa ctgacgcYtt
80761 cacaccaatg gagccagRtg ccagcctatg tggccaaatg tccttaccaag tcccaaggctg
80821 catggaaatc agagggaccc caaatcgaat gagcgagggg cccttaaaga aactcacctg
80881 gccccttgc ttcaggttccctt aggggttggt gctgccagga acttcatag gagagggct
80941 tggcccttgc tagtcccaag catctgagca gggccaaacgaa aagtgcgtt agggccaggcc
81001 tggcccttgc gaggaaaggc agagaggaag gagagggaa gcccgggttggaa ccacccctgc
81061 tcagacccaa ggcagaatcc agtgcgtatg gtttccaaata gggcccttc ttttgcgtgg
81121 ccccccatttttgc ccatgtgac cctgcacaga gctccctgtca gatctgggg gtcctccctgc
81181 actgggatttgc tcctcaatgtt gcccggccccc tcaggccatg gtttccctc tgcatttgc
81241 caagtgcctt cccagacacc ctgcagccctc acgcctctgt gaccatctat gtcacccctag
81301 cctggccttgc gttccagggtt ccctccttgc ttagctgca ctaaacatgat gagaakctcc
81361 ctcccatccca agccctgaaa tatgtcagc ctcaaaaggctt tggcacaat ctcagcttcc
81421 agatgtctgtg tgccttgc gatattaacc acactctgca tagccctctg ccatgagggt

85321 ggcacacaca gaggcctggc ccaagtgggg ttaagacaca gtgagaggag aggaatcgaa
85381 gatggcgtct ccacctaggg ggtgctgctg agagagcaa aagctggag agtaggtgac
85441 agacgaggga agccgtcagg tgtaatgcg gatagggagc agatgttgg ggtgttcaga
85501 ggacggaga tatgaaaagt catctggaa gaggcaaca gactgagggg agcagtctgg
85561 ctgctagcgg caccttgcg gtgaacatca tcgtaagcaa gacgagaagc aaggctgagc
85621 ttttgtcagg ggtaccctgc ccagcatgaa gttaggtgaa agtttcat ttaccaggact
85681 aggccattgt tcagtagatg ttacaaagca agagtgactc caatgaggaa acctgcagg
85741 taagaSaagg aggctggta aagggtggag ggaaggatca gtgggtgag agaccgtgac
85801 aaggcggcag catcaatggc ctgcagggtt gctggagcac aggtcggagc aaaagggAAC
85861 cggaaaata gaaggtcatc agaatgcac actgcaacgg attatgcagg tataggggCC
85921 gctggcaaaa gaaggtttaa ggtatgacca aggcacaaatc cagtgggag tccttttgg
85981 aaaccttgac agctgtggc atggtcgacc tcatggaaa ttttgcacaa tgatcctgt
86041 gtcaaattgag ttcccgtgac gcgtgtcagt gcgtgtggat cttgtgttt gggggaggGC
86101 agaaactctg cagacccat gcctgccRC cctcgaccct gatacgcgat gccaatagac
86161 acaactgtgc gggaaaccata aacgtccacca gattcagagc ttgagccac ctctcccaG
86221 ggtttgtcc ggttaaaggga agaggaacag ggttcaactc tcctggcaga actctatgg
86281 tgcccgagaa gacaacgta atgagtttgc cttaaacttgc aaggatttgcggYtttg
86341 aaagaggagg aaactcaat ctccaggagat ttatggca ggagtggact tgagtgttgg
86401 ttccccta ataggggtt agctgcattt gttggatgaa ccaaggccct ctgttagaata
86461 tgctggagaa caatgattca agactacaaa gaaactgaag tgtgagcga agataggagt
86521 gaaccttgcg gaaccggagg ccatgcaact gcctgtttc ggggtggagga acaaagcTT
86581 gtcactgtgt tgccaggagc agagtgcac agtgcctggaa gctcctcggg aacgttagt
86641 gtgtccctca tcgcaggca tggaccctg caccagtggc cagtacctgc tgccccagcc
86701 agcctctgtg aggacacac tggcagagat gacctcacgc tggggaaaaa cgaggagaca
86761 gcaagatcaac ccagccgggc ttccacgtcc agcaaaacgaa ggcgagttcg gttgttatt
86821 ctgctcttc atgagtcagg gcttgggccc tcactgcctt ccaaaccagag tcaaaatcaa
86881 ttgcacgttc aaaacgcacat gaaaattggc atgtcaaaagg ccattctgtt atttcagtgt
86941 attaaaatca aaagagcaca gcctcttgc agaagtttcc ttggcaga attcaagtt
87001 cctctttctt ttacttgc aatttgaatt tatttactt tttcaggaag tgaccattt
87061 tgcaggaaaaa gtgttagtcaa aagagataat gcacatgaag ttttttttta agttctgagt
87121 atgtcaaga gatttacagt gcattaaaat aaactatagg caaaatagtg tttgttaccta
87181 tctctacgca tttttaaaaa gggcaagtg ggggaccatc ttgttgcattt ttcacaccag
87241 tagccccctt ccgcgcgcgg gaagggcctt gccccaccca atgtgtctct ggtggatgg
87301 taattcatgc ccatSctact ccccaactgc tgggtgggg gggcacatag cccacaccctg
87361 gtactcaatg ccctcagcac agtattggt ccagggggtt aacacatgac tggagtgaga
87421 ctaatcatgg gtcttccatg agatctgaaat atttgggagaa aaaaatattt ctttcctggc
87481 tcagtaagct cacagcgtag gcctggggct gcttatgttgc agagccaaatc tgagaaggaa
87541 gtctctgcag aagagagccaa agtcttagga ggaggagaga cgcgtggcct tggcagccct
87601 gtgtgggcca ggggtgcgt ctgtttgttgc cgcgtgtggc agcaactYtc ttttttttgg
87661 cttaatttgc ttggatctc taatttatacc ttgaatgacc ctgacatatac cagtgtctc
87721 taccttcctt attgtgttgg cccctttaaa aaagatttggg aatMatctgS catgagaatt
87781 ccacttcctgg aatgtgttgg agaggagccc catggacaca ttccctctgg aacaaccata
87841 actgggtggaaa agtatttggaaa agtccaccatc aacaacaatt taaagtctct agaaatttac
87901 ctaagagcac acaacaaact aagcaacatc atctattcaaa gaaaatcaac caaatcttgg
87961 gaagaaaaat cccagtaaga acagcaagaa tcggggcgc ctgagccatc acttgcctc
88021 tcttccccgc ttcttagtctt tccagttcag cttatggggc gttccactct ggggggtgtt
88081 gggcaagaac acaggggctcc ctcttccctc agatcccagt caagaataac agtatttac
88141 caggatagat gggccaccatc cgtttccatc catctccaaatc tcccaactgc agagggtac
88201 tccctgctga gtgtgcggc Raggccggag cacccatc caccaggccc taactcagg
88261 caggagctcc gtctcagggtg cagcaggccg agagtgcgtt gctccataatc accctaagtc
88321 cagcctgttt gctagaaaaaa tattccatgc tagggggagg caaaccacaa agatcagagg
88381 ctaccttcctgg cccaaatccatc ggagtgggtgg cttcagtgaa gatttgcggc gcaagaattt
88441 aaggaggagc tgacagctct atgagaaacaa caacctgaaac cacagaccatc ttatattact
88501 gggaaaaaaa aaaaaatca ggagaacaga tcgttatgat gatctttcc agggggagaaa
88561 tgaatttcaaa agactggcct caaaatcatc cctacccaaa tttaatttggaa tcaaactatg
88621 tgacaattta tgtctcaggat cattttggaa aataatagag caatcatccc tcaatttagt
88681 aaggccatgt gctgggtgtt ataccaatg aaggcagagag tgtaacagac agatcagaga
88741 aaggggacagt cataggggagc cctgctgaaa tcactgtcat cccggatgac ccctgcacac
88801 gctcaaggctt gtgcctccaa acagctgggg cttcccaactg tgaagaataac agactttgt
88861 aaaacagtctt aaccaatca gggaaacaaat aaatacacaat acaactgcac acagcctcag
88921 agatggggga gggggatcag ttt
88981 acaacagaaaaa cacaatgac tcatttcacag aaaaaaaaaaaggc ggcaacagac actgctgg
89041 agaaaagccca gatgtcagat gtaacagatc tcagatggat cagcaaagat ttcaaagcag
89101 ccattggaaa tatgtcttttgc aaaaaaggaa accatgttgc ttcaaaggag tgatgtgac

89161 aatatctcat caaataaaaga ctatcaataa agagaaaatta taaacaagaaa ccaaattggaa
89221 attctggagt tgaaaagcac cataacttagg aggcttaaca gtagattggg actggcagaa
89281 gaaagaatct tgaaaacaga tcaatagaca ttatgcaatt tgaaggacac agggaaaaaa
89341 gaatgaagaa aagtaaacgg aacttaagaa aatgtggaa accataaagg gcatgaacat
89401 atgcacaatg ggagtccccag aaagaggcaa gtgagtataa aggagcagaa aagatggaaag
89461 aaacagtggc aaaagctcc ctaaatttg tgaaaacatt aacctacgtg tcggagaaac
89521 tcagcaact tcaaaccggg taaatgcca gagatccca cccagagaca tcatagtaaa
89581 aatgccggaa gccaaagaca aagaaaaat tctgaaagca gaaaagaaa aatactg
89641 cacaatgtg ggaatccccag tgggattttc agccacttc tataatggg agccagaagg
89701 cagcagaatg gcatatccaa tgtaaaaaag ctgtcaacca aagatcttc tccagagaa
89761 ctatttca aattaaagg tcaatataga cagtcccaga taagaaaaa cagagaatta
89821 tgtagtagacc catttatca gacatactaa aggagttct tcaagctgaa agcaagtgac
89881 ctcaatgtt aacttgaatt cacacccaca cacatacagt accagtaaag gtaatcacgc
89941 aactttaaa gacaatataa atgggtattt cttctcttt cttcccttaa ctggcttaaa
90001 aagcaataat atagaagaag atgtatataa ttgtatttt gggcccatat aatataaaa
90061 tgtaatataat ttgactaaca gcacaaagca ggtgaacaga aacacagcta ctggagtaaa
90121 aaaatgacac cagatggaa cttaatcca ccgcaacaaa tgaagagagc cagaatgg
90181 caacaggaag gcaactgtt gatataatg tactctcctc tctttttca gcttcttaa
90241 gcatgaatt atataaagca ataattacaa caattgttag gttttagca tatatagctg
90301 tgggttagc agaactgtgg ccctccaaag atgttcacat tcaatcccc agaacctatc
90361 tatctgtaa gttatatagt aaggaaaaat aacattgtaa atcagatgac tttaaaatat
90421 tgaaattatc ctggatttagc catggggcc cagttaaatc acaagggttca atKagggg
90481 aaggggaaag cagaagagag aacttgggg gatggctgca cgggaaaggac ttggctgt
90541 gttgctggca taaaagacga aggacagaag tcatgagcca aggaatcgag gtggctcc
90601 gaaactgaa agggcaagga aacagatct cccggagga ggacccagaa ggaacaaaac
90661 cctgccccaa ccctgattt aactcagtg gactcattt agactctgac ttccaaaact
90721 gcaaaaataat accattgcat tatttaagcc actaaatttg tggtaatttg ttacccgagc
90781 aatcaaaaac caacacaagt tgacagcaaa gaacagtaca agatggggg tggccgggc
90841 gcggtggctc acgcctgtaa tcccagcaact ttggaggccc gaggcaggcg gatcaagg
90901 tcaggagatc gagaccatcc tggctaaacac ggtggaaaccc cgtctctact aaaaatacaa
90961 aaaattagcc gggcgtggta gccccgcct gtagtcccag ctactcggg ggcgtgaggca
91021 ggagaatggc gtgaaccccg gaggcggagc ttgcgttagc cgagatcgcc gccactgcac
91081 tccagctgg ggcacagagc gagactccgt ctcaaaaaaaaa aaaaaaaaaaaa
91141 aaaaaaaaaaMaM agatgggggK ggaagaatag agctataat ggtatgtt tctatagtc
91201 acaggaatta agttatataa aatctaaaga ttctgataatc agtacatacg caaggagag
91261 gacatcacca aagtgcacaga atagggact ccaaaatttgc cccctccac aaaaacagga
91321 aataagcttag taaaatcttag aatcgacttt tttgaaactc cagaatttaa ccagaaacact
91381 acaacaacca gggaatgctt aatgaagaaaa gcagacttg aattttggg acagttgg
91441 atgggtttt aatttacttg cttaccatc tcttattctc caactcagca gtggccatga
91501 agatggcaga ccacattcct ggtgttaggt gctgatacca acctacacat ccaaggagct
91561 caatgactcc cagtttggaa actcaaaaaga aagaaatatg gaccttatat ccagagaatt
91621 gtgggttggg gtttttaaccc atctgctggc ttgcctttgt tcaactgtc tcggagctt
91681 cccaggactg aagtggtttcc ccaagacagt gttgtcaaaa gtatttaaag gtRaatgtat
91741 tagctcgaga cacctggggc acaggacaat aKttgggca agacRcacaa taaacatata
91801 taaaggctt gggagaaaga ggttagggaa ggagatattt gggaaaataa ggttttgg
91861 aagctcttgt gcataccaga gaatcttaga ttctaccat gacatgtctc ggggtgatgt
91921 catgctcaga aaaggtctaa gaagactcaa aagtttac ctcgtgtgtc catttaggt
91981 ctgtacaatg aagaaagtggaa ggctaaagaca ggttggaaa cagtctggc aagcatcaaa
92041 ggtgtctca aacacagaaac caatctacaa agacgaggag agtatttctt cctaccctgg
92101 cttttaagaa agtctctgtc aaatctactag ctgaccacta agctaacaaa acagagactt
92161 tggtgccat atataatgaa caatgcagac tttacaaaaa tagtttagaa aaatactaa
92221 acaaacaagg aacaacacat aaaaaacaac aagccccggg aagtgggggg catctgattt
92281 ccagagtctc cacattcaaa atgtccagtt tccaacaaaa aatacaaagc atgcaaaata
92341 acaagaaaagg atggtccact cacaggaaat aaaaaagacc tccctgagg aagccttagac
92401 acctcatgtg ccagacaaag attttaataa aattgtcttca aatgtgttca aagaactaaa
92461 agaaaccatg ggcaaaataac gtaagtaaaag caagaacata atgtatcaca aaatagagaa
92521 tatcaataaa gaaatagaaa ttataaaaaaag gaaccacaaa caattttgga gctgaaaagc
92581 ataataactg aaataaaaaat ttcaatagat cagttcaataa gcatgttga gcaggcaac
92641 aaagaatctg ccaaacttaca ttatcagcaaa tcatccaatc tgagaagtag aagggaaaaag
92701 aatggggaaa aaggaaacaga acctaagaga tcaagtagac caaaaatacaa aaaaatgaga
92761 gttcaagaaaa gagaatagag aaacaaaggc gcataaaagaa tttttggaaa aacaatgg
92821 ccaaattccc caaatttgc aaaaagaaca tgattctaca catccaagg gctcaatgac
92881 tccaagtttggaa aaaaactcaaa gagttacatc aagatacatt gtaatcaat tgcgggg
92941 caaagacaaa aagagatct tggaaatccac gagaaggq gcaatggc catatacaag

93001 ggctctcaat gagcttaaca actgatttg catcagaaac aataaaggcc agaaacccaac
93061 gggatgcac attaaggtg ctgaaaaaaaa tactgtcccg cggtaaaaa ttccaacaaa
93121 actatccttc aaaaatcacw gaaaaatWag acatctttag agaaacaaaaa actgagagag
93181 tttttttct agttgacactg ccctccaata agtggtaaag aaagtccctc aggttggaaat
93241 aaaaggacag aagacagtaa ctc当地agcc tatgatgaaa tgaagaatac taataaagggg
93301 aacccccaaaaa tcttcgtttg taactcttct ctttttct atatgattha gtctatagat
93361 gcgtaaagca ataattataa atctacatta atagacacac atgtatgaaa tgtaacaata
93421 ttataataaa ggggtgggaa gacagagctg tataggagca aagcattttg tatactattg
93481 acactaagtt gatatcaatt caaacaaaat tggttatgtt ttaagataat agctgttaatt
93541 cccaggataa ccaataagac catagtttt aaatatacag caaaataaat gaaaacatga
93601 atcaaaaacag tacactaaaa aacaatgtgt taaacacaaaa agaaggaaat atttgaggaa
93661 ttgaggaata aagaatatacatacatataaaa acaaatacata gatagcagaa ttcaatcc
93721 ccttatcgt aattatttaa atacagatgg actaaatgtc gcaattaaaaa ttcaagagaat
93781 gggagaatgg attaaaaaat cccaaagat aatccaacta tatgttccat cacaggggac
93841 caaatatttaa ttcaaaaatc acaaataatg taaaaatgaa aagatggata aagaaatgt
93901 atgcaatagaa aaccaggaaag agacctggg tggccataact aatatcagac aagatagact
93961 tcaagtccaa aatttttaca acaaagaaca gcattat ttatgtgaaag agtcaatcc
94021 ccaagaagat ataacaatataa taaacataaa aatagaccat ttacaataa tagctgaaga
94081 gaagaaaaaa ctgacagaat tgaagatgaa aacaaccggc acatgatcag taagaaaata
94141 cttcaatattt ccactttcaa aaatggataa aacaaccggc acatgatcag taagaaaata
94201 gaggacatga acaacactgt aaatcaacta gacctgacag acatacatag agcaccccc
94261 tcaacagcgg cagaatgtgt attttctca tttatgtgaa ttcaatccat cacacccaa
94321 tatgttaggt cacaacaaaatc gtctcaataa ctttaaagag actgaaaacca tacaaggcat
94381 cttctccaaac cacagtggaa taaaactaga aatcaatgac agaagaaacc ctttaaaaat
94441 tcacaatataatgaaatttt aacaatattt ttttaaagaa ccatttaggtc aggtatatg
94501 ttacaagatataaattt aataataat aataatctt gggatggaaatgg aaaaataaaaa
94561 cacaacatac cccaaactt aaggacatac aaaaaggatg cccggggaaatgg tatttatag
94621 ttcaatttcc tacattaaaaa ataaataaaataaatttcc tacattttccat aatccat
94681 aaggaaccag aagaaaaagca acagaaactc aagcttagca gaaaaagga aataatcaag
94741 attatagtga aggttagacaa aatagagaat tgggggggggggggggggggggggggggggg
94801 aaaaagctggg tttttggaaatc acaacaaaaa tcacaaacaaatgggggggggggggggggg
94861 taaaaagaga gaaggctcca attactaaaaa tcacaaacaaatgggggggggggggggggg
94921 accttataga aataaaaaatg attataagag aatattatgt acaattgtat acgaacaaat
94981 taaataaccc agaagaaatg gacaaattcc tagcaacaca gaaactacca aaattgactc
95041 aaggagaaaaatggaaatgg tacagatcta taacaatgtt agagcttggaa tcagtcata
95101 aaaaactccca ataaagaaaaa accttaggacc agatggctt actggaaattt ctacctatca
95161 cttaaagat taatattccat cttctatgaa gggcagcatt accctgatac cgaaaccaga
95221 caaaacatc ataagaaaaac tacagactaa tttttttttttaatataatgg gaaaaatcc
95281 tcaatgatataatgg
95341 aagtgggatt tatcggttggaaatgggggggggggggggggggggggggggggggggggggg
95401 tacactacat tcataaaaaatg ggggggggggggggggggggggggggggggggggggggg
95461 aaagcattag acagaaccca acaccctttg tgatcaaaaaa atactcaaga acaaggact
95521 agaagggaaa ttccctcagca tgataaagggg tttttttttttaatataatgg gaaaaatcc
95581 ggtcaactgt aagactggaa gcctgcttc accacctgtt ttcaacatag taccggacat
95641 tcttagtcaga gcaatttcagc aaaaaaaaaaaa taaaatataaataaaggcatcc acattgc
95701 ggaagaagta acacttcttg aggaaaaatg gggatgtat gctcggttgg acagggtttc
95761 ttt
95821 atatagtaaa aatt
95881 cagctcaata aacccacccat aaaaaaaaaaaa taaaatataatgggggggggggggggggg
95941 ctcagagccca tttccagcccc agccccaggcc ccaggcaacc ctctgtgggt gacgcagcat
96001 tacctgctt ctggctgtcc agtgggtgt agcgcaccc cctgggggggggggggggggggg
96061 gtcatgcggta aaccaagaaR gggagcaagt agtgggtgtt caaaaaatgggggggggg
96121 ctaactttcg aaacctggct gatggggggcc caatgtSacc ctctaaatcccttccat
96181 agcctcggtct ttt
96241 gcacattgtct cc当地ggactg gagccccccagc taaaacacta gacaggagag aagtccagag
96301 atccctgctca ctMtagggaa ggagaaagac ctcttcgtt ttttttttttttttttttttt
96361 aaaggccatcat gcatgggttcccttcccttcccttcccttcccttcccttcccttcccttcc
96421 gtagcagaaaa gagcccaattt ctgcatcagg cggggcttggg ttcaatccca agatcc
96481 ctt
96541 cccatggccct ttt
96601 ctgcttagaaaa cccagccagg cagggtttgtt aatgttttttttttttttttttttttt
96661 gtatttctgc ctgatgggttcaacttccat caaggacca atgttagaaat tactctgc
96721 cccaaaaaagc aggcctcaga cccccccttccat cacctccca taaacaacca
96781 gggccagtgcc tgaggggtct gcatggatgaga gtcagaaacc agcagcccccccc
caccacacag

96841 agacagggtcc ctgcttaacta gacacaagaa ttggccttct ggggtgtctc tgcgtgttct
96901 gtgtatggtg acctagcccc gcccctccag gagggacact acgtgcagga cattttggaa
96961 aacagtctt gtgtttctc aaaaagtcaa aactgtgtt atcatgttag ccagaattc
97021 tcctaKgcat atacctaaga gtaatgaaca tatatgtaca cacaaggact tgtccacaaa
97081 tggccatcgaa ggagaagctg gccacagaga agatggccca gagaggaggt gctgtgtca
97141 gaagcagctc tgggaggaaa qacgtgcaca ggccaaattc agattactga cacttttt
97201 ttaatgcaga gattgtctgc agaattacct attttgact tatctgaaa aactgaggat
97261 ctggcaacac tgaccggaa tcttcctaact agcaacaatg atcagaactc agaggtttS
97321 tgtctcaactg atttttttt ctttttctt tctttttt ttttttttga gatagggct
97381 gactttgtca cccaggctgg agtgcagtg tgcaacttca gctcaactgca acctctac
97441 cccaggctca agcaattctc ctgcctcagc ctcccaagta gctggttcca caggtgcac
97501 ccaccacgtg tagctaattt ttagatttt ttggtagaga tggggttca ccattttgcc
97561 cgggctggtc tgcgacccctt gagctcaggat gatccatccg cctcagactc caaaaatgct
97621 gggattatag gcacttgcgc acacttgcgc ttgtctcaact gatttctgtat cttatTTTGT
97681 attatctttt tccttcattt ttctttggat caatgttgc ttttcttaac tWctcacact
97741 gaaaatttat ctcttaattt ttagacattt ttcttttctt aatgtaaattt tccatccac
97801 tttgtatatt tagtgcattt attatcattt agttgtatatt ttagaagttcc actattattt
97861 ctctttgac ccatgaatta ttaaaaagta ttttttttta atttccaaat aatggggatt
97921 tttgctttgg gggctattaa ttctctctc acaggcatga tggcaggga acaagacctg
97981 ctgacactat tcttttagtt tctgtgagat ttgcacttaa ttgaccagtt gttacaagtg
98041 ttccctatgt gcttgagaaa aagaataaac ctctacttgtt cagatgcaaa attcatcaat
98101 gtcaagctta ttattttgtt tattttagaactt ttcccttatct gtgtgggtt ttttgcctac
98161 ttgatccaa accaatggaa aaaaatggt tggattctctt caaaaatgtt gataaatttg
98221 tcaatttctc tatttcaac aaaggctcagc tcaaatatgt tgaggttattt tattttagt
98281 actctacatt tagaactgtt ttatcatctt ggggaaatgtt acctttctc attaaacggt
98341 gacctgtctt tccccagcag cgctccctgtt cttaggttctt gttgagttctg agattgata
98401 agcaatgtcg gcttgcctt ggccgcctga taggtctgtt tccatactttt accttccact
98461 ttttcttgcc tttacactttt actctgggtt tagcgtccag ccagatttt aatttgggt
98521 ttgtctcgat tgcagatattt gacagctctt gtcttgaat ttgtgagttt aatccattta
98581 catttacctg gattattgtt aaatctggc ttgtttctcc catcttactc tatgtatctca
98641 atttgcctt gttttctgtt ctctttttctt ctcttataat ttgtcttata aaaaagaca
98701 tcccgccgg gcScagtggc tcaggccat aatcccagca ctggggagg ccaaggtgg
98761 cggatccctt gaggccagga gttcgagacc agcctgacca acagggtgac accccatctc
98821 tacaaaaaat acaaaaaat tgggggggtg tggcagacca tgctgtat cccagctgt
98881 agggaggctg aggccaggaga atcaactggaa ccccgaggc agaggctgca gtgagccgag
98941 atcgagccac tgcactccag cctgggttac aagtggagac tccatctca aaaaaaaaaa
99001 ggcattccctt tttctctaat ttgtcctgtc accactgtgc tggtgtcccc agccaggctg
99061 caaccctggg gacacacatg cccctccagc aqgtgatcttca gacccttctc acccttggc
99121 ctcaggctcc tcacttcttgc cccctctcc tgagagggtt cacgcccattt ccacacca
99181 tccaaatctg atcatcggtt ccagtccagt ccctgcacca gctgtctat ggcctttctc
99241 tctgggggtgg ggctggggac tggatgttctt gctgtgtgtt cctgcaactac tccacatttc
99301 ccagttcaga ggtgatcgga gcccctcttca atatatgcca tcccaaggcgg gggcagcagg
99361 gcaacaaggcc acttcagttt caacacggat ctcactgcag gaactaggag gggctccct
99421 taggaagagg cactctggac tggatgttacc acagtccccca ctctccctc ctgtctcaca
99481 gttttcgatctt tccatcacac acgtgtgttcc tttttttttttt attatgttattt
99541 tatattcatg tgcgtggca tatgtatgtc ccatgacaag gggcccttgg gcttttagat
99601 caatccttag aaggacttca gccagacccc aaaggccaca tattatcaat tgggtttctt
99661 tgaagtatac agaataaggca aatccaaaga cacaggagttt agatgttgc tcaactatgag
99721 ctgggaaggd gaatggggag tggatgttctt ggggtgcaggg ttcccactgg ggtgatgaaa
99781 acattttggaa actagatgtt gatgtatgtt gcacaacact gccaaaaacc actggattac
99841 acactttcat actcatatgtt attatgttcc attgatttt aaaaggctttag aggaagcaac
99901 tcaaattgtt tggatgttctt gaaagaggcc cagatgggtt aagtgggggg ggcgtgcasa
99961 ttaagcccac aagaatatttcttcaacttccatc caggttcacac acacactcYc tctcaactgt
100021 cctgaaggcc atagccccctt ctgcctgtact ctccatcttca ctgtKgatgg cgtcaaaagat
100081 gcaaaaaat atattttggc tggatgttacc acaacaggcc acaatgttggg agggtggcca
100141 agcaggcttat tagcaggatt tggatgttcc tttttttttt ccccccaggca
100201 gactcttgc ggacagagac cagattgttcc atactcgtt cccaggcaggc cgctggct
100261 ggggtgttgc agacactgac tggatgttcc gatggatgtt ggggtgttggat gatgttgg
100321 gtagatgtt gatgttggat tggatgttcc gatggatgtt gatgttggat tggatgttgg
100381 aataattggaa tggatgttggat ggttagatggg ggtatgggtt gatgttggat aatgttgg
100441 atggagatca atgttggat gatgttggat tggatgttcc gatgttggat gatgttggat
100501 cacgtggat gatgttggat tggatgttggat gatgttggat gatgttggat gatgttggat
100561 tggatgttggat gatgttggat tggatgttggat gatgttggat gatgttggat gatgttggat
100621 ctggaaaggtaatccatcacttccatggca aatgttggat gatgttggat gatgttggat

100681 accatagcag cgtcatttgc cactgttgcc atttatacac attttgggg cagaatcggt
100741 catgcttgcac cacaggcat ccttagaggg tgaggattgg gggcaggaa ctgtgggtg
100801 ggctcagggc tagggctgcc gatgtggcct ttgtgcctgt gaggcactg catttccatc
100861 agcctcatcc ccatcttccc tctctgcctc caccagtgag ttcaagatcct tcctacaatt
100921 caatcttggg gtccccctac ttctgtccc atctcccaag acctggctgg gcctgactc
100981 agaatgggg cagcaaagcc ctgaggtaga gccgtctccc aagattatgg cctggctgg
101041 gcccgcctca cctccattca aaggacacca tggaaaac caatgtgcc ccaagcaggg
101101 gtggcacggc acagacactg atggcacagg ggcacttct tgctccttcc tgggtgttcc
101161 atcctcgggg atcaggggct ctgcagcagc aggcaaggct gggcagccct ggagtgttg
101221 acacaattcg ggcactgtga ccctactgga ccccttgca cagatgtgS tgggcMtagg
101281 aactggctgc tccccactgg ctgctctggg aggccctggg gtcgtcctac acccccaaca
101341 catgagcatg ggcaggattc caccagtctc YttctStgtc cccagcctct cccgcagctt
101401 gccttggtc caccgtggac agcgtctgt acacggcYt tcttccaca gggagacccg
101461 gagggcttgc ttctgcctc agagcagagg ccccaggctc cttcagggtgg gagcacaggt
101521 gccttcgcga cagggccca cagccagctc agcccacaga acagtaagcc gccaggtccc
101581 aagttgtcat cttgtgaccc gcccctgggg acacgggtgc ccaacaggaa agaggacggg
101641 ggcttattat aaggaaaata gtgtcaggcc tgccatggag ctttgggcct gaggcctcag
101701 ctgaggccaa ggcagcagga ctcatccaga gacaggacag accctgggcct cctgggtccc
101761 agctcaggga agctcctggg tgcggctcgg ttcactccag catttccatc atggatgaaR
101821 atgtggatgt gtctgtctg accatgggtg tgtcatgcc tgcgtgttg tgagagagag
101881 agaggagaga gagacggaga gagtttatgt acgtgttaact accaaccaga gtccaggagt
101941 ctcacactcc actgcaaacc aggccaggcgg ggcacatcg agtggggggg ccaggcacga
102001 atcggcagggg agggggggg tgggttcggc catccagagg aggacacag cctccaccc
102061 agtgtctgtc gcggaaacac gcccagggtg gccaggctt ctgaaggaac tgaaaaattga
102121 gatttcatg tgaaacctcc tgattttaa tgatcacatc aacatttctg aagccgttgt
102181 acgagtcaaa taaaacattt cgggggtgggg cggcgaggcc gccaagagc tgcagggttg
102241 aggctggct ttatgtccagc accccgatgt acatacagac gagaagactg aggtcaggcc
102301 agagcaagga ctccgcaggc caggcccaca ctggcgccc cgctgcctg gcccacact
102361 ggggtcaca cagccccatgg ggagcaccaa gSaccacaag aaaggaggg ccaggcctgg
102421 tccccggggt ctctgtcacc aaggggccac attgtgagtt tcaaagggtgg cggagggtgg
102481 tcctacaggc ctgtgagttcc ccacacttcc tcccagttca aggctcggg acaaattccaa
102541 gcccagctt gtcctcggt tgatgaaggc caagaatgtc gggagctccc gggatattgc
102601 cYaKaAaaa ggagcacaagg cacctgacgt tctggacaca agggcagaac cctttgaga
102661 ccatcatcg cctttgtc tccaggctgc gggggctgg aggacatcca ggcacggta
102721 taggtgcgg gagccccggg aacagaaccc tggggccga actgggcacc agatgcacca
102781 gcccgggggtc tgcggccca gcccggggag tccagggtct gcttccgaacc cggcgcacag
102841 cctgctccct cagtgcggca tccaacagct ctgtctgcct gggctgggc cctgcRccct
102901 ccccgctgct ggctcccgct gatgcgtac agatgggtt catcgctgcc tccaagaaca
102961 acaggacgac cgcagggttc gcctcacaga aacggaaaccg ctggaaaggc ctctcttgc
103021 attccccaaa ggaaaataaa aatttctctg aagtacattt caactccgag gcagccctgg
103081 ccacatccaa gaagctcggt cctgagcgcc gtgggctcca gtgaagggtcc ccagcaactga
103141 ggccgcctcc tgaccccat cccggccact gccccgggtc gctaaccct tacacggctt
103201 tcacaaaattt tgaatttagt gcaacagtt taaaatgaag agattgtaa taaaatcaat
103261 ttccagttc tcttataaaa ttggaggccc tcctgtctg ggccgacccc gtgtctaccc
103321 gggcccagtg atctctgtc gcgtcacccgt ccaggcccccc atcagtgtgc ccatacgaga
103381 ccactgtcct cagcgaacgc ctgacccccc ggtccaggc aagggtgaga ggacagatgc
103441 ttcatctgtc ggggctctgg cttctctgg gccagtttca tccagctctg ccctccagag
103501 ccctctgcac acgctggggc tggactgtcc catggctgtg ggctctgatt tgggtggca
103561 gacctgagca ggcacatcatg cttcttattt ccagcctgac accccccagac cggccccccgc
103621 caacccagac aatgtgtact gaactctctg tcaacttgc ccctgagctc ctcccccgcag
103681 cccaccagggt gtcctgtc tgggttggc agcacctagt cctttcgcc ctgctggacg
103741 tcgcctggcc agacttgca agccatctca gccccttgc tcctgtactc catccctcca
103801 acgctggcca ccccgccctc ggcggccagct cctgcccacgt gaattgtcaa aggcagagca
103861 tgaggaggct gaaggcagcc cccatggcac accccacggta gtcctgtcag tcgtccaaa
103921 aaggctcaag cctggcagc attgtctgc gcttagaattc tgcagctgga ctcaaatatg
103981 tgcggcaat gatctggaa aggtcttgc gaaaaggccc agacacaggg accttccag
104041 gcccggatcc cacagaggaa atgattctag cgagcgcac ctcgacaggt ctgaggtcca
104101 gacaggggtg gagctgtcac aggccaggcgg gtggggctgt gatctggcct ctcccagcc
104161 cacggtgatt tggaaatgggaa atacggttct gccaggcaag gagtatcccc ctctgacc
104221 cctgtggact gcctaccagg actcccaaca cccttcctgc acgtggcgc tgagggctg
104281 atccggccgc tgctctcatt cccatggcca cctgatccag gccgcagctc ctccctgcac
104341 acaggacagc agcctcctgc cggctccctg ccttccctgt gctcaccctc ctccggaaacc
104401 agaccatgccc acctgtcga tggcacagcc cctcctccac gacggagccc acagaagcc
104461 tcgtatctg acctccagtc ccctcaggcc ctcaccctgc ccctcctgc acaccccaaa

104521 acaagccccg tcctcacccc tcaggtactt cgcacatcgct gccctggccc cttttcctcc
104581 tgggaaactg caacccacac tccaggcctc agcccaacgt ccagtactg cgggctca
104641 ctgtggcaact tcccaagtct ttcatgtgc aggtatttg ggactgacta gtgactgtg
104701 ggctgtatga ggtcaggcgt gtgcctgtct ggtcccagag cttggccag tgccccgtcc
104761 tgcgtgcgt gcccggtaaa gatctgttg atgaatcaag ggagagagga aggggagcag
104821 ccctgggatt ccacagggtgc tttatgtcca gagacaaccg cgacagggaa agtggccaag
104881 gcacgctggg agcagtgtc cagccgacg qcatccccag ggcagaggct gggattcacg
104941 cattgctgcc ctcagggcacc aaacgggcct gtggcagtca caggggcaca ttccctgata
105001 cgggtctgtc gttccagcat ggacaggcct atttccctcg ccatctgcag tctggcatca
105061 gacattccca aataaaggac agacagttaa ctgtcagaaa gtgcgtgtgc ttggccctgg
105121 aaggaaggaa gcccgtgtct actgggcacc cccaattttg gccactgtgc agtgcctgt
105181 gtccctcaacc tcggggcctg gaccacagct ccatgagacc acagggagga agcagagtc
105241 cagaagaaag agtgtatgaac agcaactggcc acttgatttg cagggcacag tgcaaaatgg
105301 aaacgcacag tcccttgttc aaaaagtgtatcacttct aaggcagggaa tgacagagtg
105361 ttaaagtggg cataggccc tgccggatgg gtcacaagcc ggcaagtca gtccctgctca
105421 agtgtgagga agtgggacta gcaaaggccc aggggccaccc aatgtcccat gggccatgg
105481 gaccagtagc caccacacgc tgccaaacctg agtctgcagc cctgacactg tggcaacaac
105541 aggaggcgtc agtgccaaacg cctccaaatgc ctgcagggag caactgttgc tcccggagtg
105601 agctcctggg cgaccggcgt tcctaggcac tatgcttatt ctttctttc acggatattt
105661 acgcaccaac cgtgtgccag gcattcctca aaatgctgca aattccactg taagaaaaag
105721 attctccccg aagctaactg gctaagtgc acgtggagacac atccctactc accctcgctg
105781 gggcccaggct tctgcagccg ttaagtggag ttccaggcct tgggtgttgg aaggccctct
105841 gtcccaacaaa cagatggggt cacaggctac cgggttttgc cagcccaggg tctgagcggg
105901 gacaggaggg cactgcactg ggtgtcagac gcccctgggtt ctaacgcctgc tccctccctcc
105961 ttagtgcagga gtcaaaaaggc ctggctccca ctgttttac agtactat ctgtgcaacc
106021 gtcggccacct cagtttctc atctgtaaacg tgggcacagc aaaagcacca acccaccgtg
106081 gggctcatta caagaatcct aggagatgtatc gtaacctgcag cgtgctggca gacgcggc
106141 acggtaagtgc ctccaggaac ggcgtcacgtt ctttctgtc cctgctgttc cgcgtactgc
106201 ttggacagtg gttagccctgc accttagttt tctctctgtc tccaggacta cgcgtccat
106261 taaatgacac catgaagagt gcccacccctc ggcgcaggccc tcggtgaatg ggcacatgt
106321 tgaacgcacaa cttggctcca acggggcagg agaaggatca ttcagactca gctacagaaaa
106381 aacagacttt tgcgtcaacta acgcttaacaa aactggaaag gaaagaggaa ctacggcctg
106441 gggaaaagct tgcgtcaac atgacagaaaa gtcataatct aaaaagagat aaaggccta
106501 tc当地atata taagacccca ggagataaaac aggcaaaaggc ctttagaataa ggagacatgc
106561 agagtaatca aaccgggaa agatcactca tccctgcacg taaccaaaa agcgcacaaatt
106621 agcgcacaaaa gcaacgcgcg tttgctgccc gaggcggaaa atattaaaca ctaaagcccc
106681 cgtgctgggtg aggatttggaa gaaaccaggaa gctgctgtgg gagccagttt ggctctgccc
106741 ctagtgggtga gtctgttaag cgctgggatt tcctaattca ttttagtcaact cagtaggtgc
106801 cccctgatgg ctgctcaggg ccaggcgtcg agcgcgtctg ggctggctca ggcgcaccc
106861 tgccgggtca gcaacagata cctatttgc ggaacagaac ctgaggtcag ctcccaggat
106921 gtc当地caccct ctgagttca agagatgtt caccaggagc agggagcatt ccagctgtct
106981 cgggcagctt ttgtgtggg aacagctggt agcgcgtcct gcccacccctt ggggtggaca
107041 ggaaggctct gcctgcaaaag gcctagaggc agccacacac cctgcccgtc acactcctaa
107101 gtacgtgcag gccacagttc ctgcacatccat tc当地cccg cagtctggg tccctcccca
107161 ctagtggcagg caaggggaag ctggaccggg cctgctgatc acacaaggat gcacacacgt
107221 ccatgggacc ccaaaactca cactgtccct gacagccgc cccacagagc acagctctg
107281 cctccggag caagaggagg tggcagtgc tccctgggccc tgggtgcctg tggggctgca
107341 ttgggggctc ctgccccgc cccgtgagtc tgggaacaca tgggtcttgg cttggcacc
107401 ttctttctg gaaaaacacc tccctctggg ccaggcgtcg gcaagacactt gcaacttgc
107461 agctgtggag tcagagtgtatc ggggtccaaat tctgacttcc gcaacttgc cagaccac
107521 aacctgcccac ccccaccaggc acacctgccc gcccacccccc accacacac ctgccccggcc
107581 agcctcaggc caccaggccag cctcccgaccc acgtgaccc tcaggggcag gatgtgtc
107641 tgattccctc ctccacacca cgtcccgaa gctccaccc tc当地ccatcc tgaacccact
107701 gc当地ccaccc ccatctgtca ccgggactcc tccctgtcc caccggcgtcc accgaaaacc
107761 ctcccaaggc ctccacccgc caggaaatgaa atctccaccc cgtaccgtga ccgtcccttc
107821 aaagcgtatga gtgcacccca tccctcccg ecctgcacgc ctc当地ccactt cccattact
107881 cctcagttcc aggtacttgc tcttcccttcc gccccttggt caccgtgcgg taactccagc
107941 ctcaggcgtc ctgcacatgtc tggcccttc gcttggaaaca cccttcccttcc ggcgtggc
108001 tcccatcatc accaacctgg ggcccttccc tcccttaccc gccttctcact gctgccc
108061 caccgggtt tactttgttcc acaggatttcc cactgtctg cgacgatctg ctcaccagg
108121 ctccctgctg gggctgtgag cccagagaac agggacttgc tgggttcct ccgtgt
108181 agcccccagca cccagttctg tcagacacat ggtggacaca tagcaagaac gtgacgagtg
108241 aatgcaggaa aatacacttgc tggcagcagg caaggtggcc tcttgcact tcagtctccc
108301 cctaaggactt agagctgacta atgttattat tcccccattga tttggggagc agtattattc

108361 ccccatgatt tggggagcag tattattccc ccatgatttg gggagcagca ttatcccccc
108421 atgatttggg gaggcaggatt attccccat gatttggggg gcagcattat tccatgattt
108481 ggggagcagc attattccc catgatttgg ggagcagcat tattcccca tgatttgggg
108541 agcagcatta ttccccatg atttggggag cagcattatt ccccatgat ttggggagca
108601 gcattattcc cccatgattt ggggagcagc attattccc catgatttgg ggagcagcat
108661 tattcccca tgatttggg agcagtatta ttccccatg atttggggag cagcattatt
108721 ccatgattt gggagcagca ttatcccccc atgatttggg gagcagcattt attattgagc
108781 attaatgaac ccgcattggca tttaggtgcc cgccctttc ccaaccagac cagcgaccca
108841 gccacccttc cacacaggcc cagtgccca gtcacccct gtccgtctc gcaggctctc
108901 tctgggtgtg ccgacagttt gggatggggag gcgtggggagg cgtgagctgc gctcctggcc
108961 gggttctcta ccagctgtgc tcgtggcga gcccacacct cacggcaagc cccagccaaa
109021 ggaccaagt caagacgtct gagagacatc ccagccgaag aatccagccc ttgtcccaaa
109081 cgtgttatg cacaggcaca cgaagaagg tgagacaga ggagagaaaa atgaagccag
109141 agacaggag agagagaaga cagaccccgac cacgcccatt tgactgcgc acctttcccc
109201 gctcagagct ccttcaggc ccagagcaca ggcgtccca ggattaattt caaagagaaa
109261 cgggtcaaag aaaatttggaa attttatag aaaatatttgc aggttggag ctggctgaag
109321 tgcgtccaa cctgcagaga attctgagag gagatggggc ctgtggtag aggcctccgg
109381 ctggcacaaa ctggcagagt ttggacctcg gacccaaatg cagttctga gccaggctg
109441 tcctcgctgc tgctggggaa gagttcctcg tcccttcag atgtccctac agtcgcctc
109501 gcagccctga tgctggtaacc caacgcctt caccgtcgc cccacccctc tctatcccc
109561 agaggcctct gtctaggac ccactatggg atgcagtccc ttgtcccca ttatgcaggg
109621 tgcctatgtc tctttcacccg tcatccatgg ggcagccagg gcatgaagga agcctcacct
109681 gggtcccaag caggtccca tgcgtccctgc ccaaaactca cccaccccgag caggctctca
109741 catccaaaggc acagatgtca cagacatgaa aagatgcctt ggacacatgc ctctctccca
109801 gaaggcgtcc tctgccatgc tggggacccctg aaaagagggg gcatgcggc cccacagtgg
109861 ctgatcaaac gccagatgcc cccgtgagca gatcacagtc tcacccctga gccgtgtact
109921 gctcgtatgc cagggactgt taccttgc aaaaacccgaa aagcttttgc gaggacattc
109981 ccaacacaga tccaaggggaa gaaaagaaaag tatcaattca ctgtgtctgt ggaagatggc
110041 gaggcagagg agggaaaggg gacggagagc aagtgtcgcgt gtgagactg ggcaggcgcc
110101 ctgtccctgc cgtaatccca gcagccctgc tgccactgca cccaggtaca gttggagatct
110161 cagcaccctcg gttctgagat cctgaggggc tgggtgagct gctctggcc ccttagtgag
110221 ccagaacaaac agcatgtcgg gtggagctga gcacactgac ccctgcacca acagcagcat
110281 gctggcgga gctgagcgcg ctgaccctgc caccagcgc gacccctgc ccaacacagg
110341 catgcggggc ggggctgagc acactgaccc ctgcaccaac agcagcatgc cgggtggagc
110401 tgagcacact gacccctgc ccaacacaggc catgcccggg gggctgagc acactgaacc
110461 ctgcacccaa ctgcccctgc cccacccctc tcacccacctt agtgcgtggaa gcatctgtgg
110521 gagcagccct aatgcagctc ctcagaacca ctggccgtc ttccagaact gtcacccctgc
110581 ctgcagttgg ggctgcagcc aagccagggc caaggccacc atcctccggg tcaggcttgt
110641 cacctgcccc cgggtctcca ccaaagatcc aggccttc cagccacatt tacctggctc
110701 gctccacccc tgaaccagaa accatgtcgg gaggggagag gtggcggggg aaggccaagc
110761 ctgagactca gggactaaat ttatgtcgg ggcacagggg agttgtcgcg cacgtgggg
110821 ctgtcaggcg gagtgcaggccc cccagacgcgt gtggaggggg agcacaccct ggcacgtctc
110881 cactgtggaa taagggggc ccagccctgc ctggggctgc aagcctcgct tctgccttca
110941 cagcccgagg ggtggggagg acccacaggc cagcgtctgc agcagcagac tccagccaga
111001 gagccagcca gggccgggg ctggccctc ctgcctctcc cgggtgaccc caggcaactc
111061 cagcctcgga gaatgtcgg cctcggtt cccagcgcg gccccggccc accccttccg
111121 ctggggccac atcaccgcgc tggcatctcc aacagccatc ccaaggggccc ggtctgcacg
111181 aggacaggag cggccggcg tctgccaagg gtgcgtctc cccctgaagca cgcctcccg
111241 accgcctcg ggtttcacaa tcagtgcctc tccctgggtt tcccgggggc tgggctgcag
111301 gtcagctcc agggtgaaa gagaggatga gctgagagga ctccaggat ccagcggcc
111361 ggcgctgaac tcagtgcgtc ggcaacagtg gcttggcagg gcccaggctg accagggtat
111421 gagagagac acctatcacc acgcccaggcc cctggcctc aggggtctgg ggagtcaatt
111481 gggaaagtgc cgggtggaga cacagagaca ggtctggca caagacgtcg gccaggctca
111541 gggctcaccct tgccaaatcca gccaaagccc tctctccatt ctccctttagg acccaagggt
111601 caacacttag acatgtcgc tgaatcatga ccccccggacg tccatgtcctt aatccccca
111661 aactcgttag cattactta tatggcggaa gagacttgc aaatgtgttt aaatcagggg
111721 tcttggatg gggaaagtaa cctggatgtt cccagcgggc cttaaatgtt atcacaagt
111781 tccttctaa agggagagct ggctatagaa gaggagaagc cacgtgacca gggagggcaga
111841 gactgcagca atgttgcac aagccaaggaa acgccaacag ccccccaggag ctgggggg
111901 ccaggaatgg attctgcctt ggggctccc gaaggaacca gcccgtccag caccttgc
111961 ttagccctta aaatccatt tcagagtctt gagctcccgactgtgagga atgcagggt
112021 gctgtttaa gccattaagg tgggtgtac ctggtacagc agccacagga agctgagatg
112081 gacgtcgaca gctacaggc agaaaaggctc tctgacccctt cattcagcaa attggcaccc
112141 cccaaacaaag cacacggccc tcctggcccc cccagccccc acctggccca

112201 agcctcacat ccgctgtgtc agtggggctg gctatctgtc cgccactgaa tggatgagtc
112261 tcgcctgctc cccagtccctg tggctctgaca gaacagagct cctctctc tcctgtgcc
112321 ttgctccagc ctccccacc acgcctccctc acaactgcaac cagtacgagc ttgctgaac
112381 actcaactct ttgtgagact ggagggcgcc atccccacct tgggtcattc aaggctgctg
112441 cttccatgga ggctgccccca cagctctgtc cctccagccc tggctctgc acacccagg
112501 ggaactgtca tgccctgaggt ccgcggccac ccccgccgccc cactcttccc ccaggatccc
112561 ctgcggcctg gaaaactctc acccaactca aggcaaaagtc acagggtgga gactgtgggt
112621 caggaggggc tctctctggc ctggtgggga agggctggct gtggctgca ggtgcagcc
112681 acgctcagcc acatggagga getggccatg gccggcgctt gtectctgtc teetggaggg
112741 cttcccttgg tgcagcgagg ctgagagct cccgtggg cgtgcctctg ctgccccgtca
112801 gggaggatgc tgcaagctgaa cagcagcagg cagaggcttgc gcaccggcgg tgtctgggt
112861 ggaggcagcc aaggtcaactg gcacgttgtc gtctgacttc tcttccttca gcccacagg
112921 gtgtgtgac agcagcctgg tagcagtttc ccagggatggg aagcacctgc ccctccctg
112981 tcccaaggcc agggccggggc gagccccagg agagcagccc acagaggagg gcagaagatg
113041 cttctacctc tgcatacacag aaagcatgg tagccaatgg gggaccgtgt tttgtttgt
113101 ctatatagtg agcacgggg agggaccctag gcatgacagc tcccactgg tgtcagggc
113161 ccagggctgg aggagagag ctgatgtga ggcttctgtg gaaagaggc tttgtctgga
113221 ttgtctggct gaggccctgga cctggccagc gtccttgc caggacctgt ctctgtgtc
113281 ccaaccaggc acttctctgag tgactccca gacccttgag gcctggctct ttegcatgtt
113341 gatagggct gtcctctgtc cccgggttgg cctggggccct gccaaggccag ctgttgcga
113401 ctcactgagc tcagcggct ggctggattc ctggagccctt ctcaatccat cctctcattc
113461 cagaatttt aaggcacat ttaaatccatc gaaaggatc aaaaaaaaaatc tgatgtctgg
113521 attttcttga tagatgttc agctgacca gacccaaactt ctaccaaattt gaggtcaagg
113581 gagttgagga gggctgttcc ctccttgag ggcctccca gacccctgtg cccatggcaa
113641 ggccagggtc ggacaccacg tggctgtca ttcccactgc tttttcttt tagtggagat
113701 gtattcttgg tctccacgtt tccataaaa ggagagaaat gaaagccagg ccaagggttgg
113761 gtgcaattat ctttttttcc tccactgtc tcctgttgg cattgcctgc cagggctc
113821 aggctctgag ttgcagctc ttgtcttttgc gcaaattggca ctgtgtgtc gaaacagaga
113881 ccacatgctc aacacgtgac tgggggttct tctcaactgtg tggcaacctt cacagagtcc
113941 caaggtcaagt ctctggaaag ctaagacagg ccctcaatgtc cagcacaggg ttggcttctc
114001 tggcacgtg gtctggagt cagccgtc aggtccctta acgcggggcgg tgggggagga
114061 ctgcacggct gagactggca cttagacatc cagggaaagag ctccacacg acagggtgg
114121 gggtccatg agggaggccag caggaccagg gcctgcctt gctgtccctc tgggttatat
114181 gcccctgggg aggcttaatgt aacgcctacatc tgctgtctt ccgcgtgacc cagacggcg
114241 cctccgggt gtgtgtgca ctgcgtgccc agccttgc acgtggggcata ttccatctt
114301 gtagctggct gctgggtgcc cccaaatgt ctggatctg atccctgcaaa cctgagacgg
114361 gtaccttca tggaaagaaa gggctttgca gatgtgttgc cagtaaggctt cttgagatg
114421 gaagcctgtc ctggattact cgggtgagtc ctaaatgtaa tcaaggcatc cttgcaagag
114481 ggaagcacag ggacacgtca gaagaggagg gggcgcacatg atccggggc agagattggc
114541 atgatgcagc cacaaggccaa gggaggctgg cagccccagg gagcggggat gggccagaaa
114601 cagatttcc cctggagccct ctggagggag cacagccctg tccaccctt atcttggctc
114661 agtgagattc atttcagact tctggccac agactgtga gagaggcatt tttttagatc
114721 ctgtttcaca atgctgataa agacataccc gagactggc aacttacaaa agaaagagg
114781 ttatggact tacagttcca ctgtggctgg gaggcctcac atcatggcag aagatgtaaag
114841 gcacgtctca catggcagca gcccggagag agagacttgc tgcaaggccaa cttccgtt
114901 taaaaccctc agatctcgta agacccactc accatcatgaa gaacacgtatc gaaaagacc
114961 acccccattga ttcaatcatc tctcaactggg tccctcccat aacacatggg aattatggg
115021 gctacaagat gagattcagg tggagacaca gagacaaacc atatcaatgt tcctattgtt
115081 tcaagctccg tgaggactcc ctctcaaca gtcccacaca ggtatgagatg cacacacacc
115141 agagaggaca atgccacagg gtgcctgtc cctgtacccc agcacagtgt ccctggacac
115201 agagctggcc ttggaaaggc cagagaacag ggattggaga tgcctgtca aatgcctgg
115261 gggacaccctc aaccactgtc ccccaaaagg ccaccaggcc tcagagctcg gcctccggc
115321 ctggccagggt tccaggccca gcagactga gcaggaccct tctgccaagg ccagcatggc
115381 cagggaggag agaaaaggac cctcccccgt ccctcaccccc ttggccctcc taataacatc
115441 ctctcccccggg tggctgtctaa gctggccagt gcacacactc atccctggcc tgcctgtgaa
115501 atgctttaaa accgtcattt tctatccgtt tccccccaggc aatctgcaat ccagaactg
115561 atcttggat acagagacta ttatggaaa tgatcttgc aggtccatg agtcccccgg
115621 atcaccctgaa gcagcccgtg ggcttctgg gaaaaacagg acatctgtaa aaggaggcac
115681 ccctgacgct gcagggatata cctgcccacc agacagatgc ccaccccccctt cttggggac
115741 cagggcagggt gtgtcacctt cccccggcaga acataaaatag caggggtgtt cacacccact
115801 tccagacgtt gctttgggg aggcaggaaag aaactgcccag ccagcaccc tcgtccagctt
115861 ctgcccaggc cacaactgata tatcagggtt ccctcagaat tcaactggaa aaggaggccc
115921 ctatacacac acacacatac acacacaccc gcacccagggt cctggaggtt tgcctggcagg
115981 aggctgtgcc cctcaatgt cctggccaca ggacacagct tttatgtatc tcccaatgtc

116041 cactctgggc cctgaaaagg cccaggcaga tggaaagctg gacgccttag gtcaccgtgg
116101 gcccctccacc cagctccggc actcaactggg cctgggtgccc ttccagtcag tgggtcgatc
116161 tcagccccagg gccacgtggg aacgattagg tgcatacggttgggactgc tcactgaagg
116221 cccgagcccc tcatacggttggg aacgattagg tgcatacggttgggactgc tcactgaagg
116281 gccccccagcc ccaagttctg ggatccagac ctatagcccc agggccacgc tcacatccac
116341 agcctgggcc ccaggcctgg agctaaggct ggacaggctc acctccccc cgacagccca
116401 ttaccaaacc cccacccccc agccaggcct cacccttagtgc ctggcagctg gcctcaccca
116461 ggcgtttctta aggccaatgg tttctgggttattctgaga ttggaaattcc tttctgtgt
116521 tctgaatttc aagttctcat taaagatatg tggattgggg gtgggggtgg agggaggatt
116581 cagcagaaga tgccaaacac cagcagaagc ccagaaaata aggaaggcc ggccaggctg
116641 tggggaaagg gctgcatgt tcaggaagac aagcaccata tggaaatctgg gatcagagac
116701 atggccaagg gcacgtggc agtcacccctc acctgagcag aggccccagg aaccaccctg
116761 aatcttcagg gagaggccag ccagaagcac caagccaca gcacaagtga ccagccctg
116821 agctctcggttggc gacacaggggg ttgctaattgc cactgggttggc gcccacgc agctagctg
116881 caaggagcaa agctggggcc ctgccccccgg gtctggagca ctgcgtcagc cccttccacg
116941 ctggccctcc tagattactc ctcccttcctt ccctgaaaaa tagtccaagg gtggggataa
117001 gtcgaggctg agaccacagc tggaaaaggc tggcaggca gtacccagtg gcccggctgt
117061 ctgcagcata gcagacgctg atttgtcacc aaggaacaag gggccaccgg ttgcactgat
117121 tggctccaga gtacattggg ctccctgtatc gaacaattcc tttgcaaaagg tatctgatgc
117181 cgtttcatga aattagaagt atgcgtgcca ggctgtaccc aattttctg tcaagctagg
117241 aagttagcaca gacctacagt aaatgaggcg gggctgtat ggagtgcgg gagtaccta
117301 gtaaatcaca gcacggggcc tggcccgacc caggccccctg gggactcgc tgcacgtg
117361 cagggtgtgcc cgatccctggg ccccgacact cagactctcc tcacccagg gtcctgacca
117421 gaggctcgagg gctgagctc cagtttccca ctctcccaagg aagggccaa cacaggccag
117481 gagttccctga ccacacgggg gaccctgaga actatggcc tgcaaaacca catcagacac
117541 caaccctggc ttcaccacaa tccaggacca gagactgggg gatgtggga ggtgtggct
117601 ggcacagtga ggccagatct gaggcgacag tcatggactg tcacccatggc cgcctggcag
117661 ctcccttcac gtttctggc ccaggccaca cgtgggttct gagttaaaaa agacagacgc
117721 ccatgagtct gtgcagagac tgctgaggct gctggagca cttccagagt cttttaggg
117781 cacatgcctc cagggatcca gaaagagcat gcaggccct tctcccttag cagccagggt
117841 agggttggga cttaccacag gcacccctccc agggaccatg gcagcacaga gcaggagagg
117901 cgcgtctgg ggcgtctggg ggccttcac accgcacacatc tcaagacgc aggcactgga
117961 ctcaggccctg gccccggggcc cccggccccc tccacgcac ttgttgcct ttaggcaaca
118021 gaagatgggc cttggacacatgttccatggc ccacatggc gggcccaagac atgaaggcag
118081 aggtgttgg gctgctgtc accccatgtc accccaaaca gaggccccag ccccaagcag
118141 cctgctacgc cctgcccgtca gccaggcagc tggcccat tctgagccact gagaccgggg
118201 ctcagcctca gcagccaaac tggctgtgg agagctgcct gagttctgg gatgctcccc
118261 tggggaccag ggcctcagag ggctctccac acagcgtc caccctgcgt agcctcacct
118321 gtgtccaggc cctgtggccc cccgagagga ggggacgtgg gtggccggcg atacaccaca
118381 tgcacacggc cttgctcagc ctccctgcgc gccagccct tctccaaggg ttgcgtgaag
118441 aactcccttcc cctccatccg gatcagacca gcctgcggg caaagacaac aggtcagat
118501 ttccagcaca caaaagacca aggaaggggg acagcgacca gacccctca gggagtacct
118561 cccacctgaa gaccctggc agcaccggc ctgggtgtca gctgtggacc tgggtggcc
118621 ggagggtatg aacatgtgtt cccagaggcc gccccccccc gcccaccctg tcttagggaaag
118681 gggcctctgt ctgggtcatc tccacccttataccaacc ctgtgtgttt tcagaaaaagt
118741 tactgtgttt ataaagtcaa aaccgtccag agaaggcagg agacattcaa gggcagctgt
118801 gggcctgtgc gtgtcccccag gagggtatcc accaaagtgc atcaatatgg gccaagtgt
118861 gccccccactc ccaggagccc agagtccaaaccaaaaggtaa acacagccag ccggggattg
118921 aatctgcccc cacacatgtt ccacatggc caaacaatattttaaagct ctgaagctaa
118981 cactcagaat gatgagattt catgtcttgc acatggaaatc ccagcttcc ttggaaacgc
119041 agccctgac cacagccggc cttggatggc acacatggcc ctgcacttgg ctccagttccc
119101 tactgtcccccaactgtctc tcctgtctc tggcgtcacc accagccccag actctacagg
119161 cagctgaact gctgcccctg acccccttagg ccagccagca ccctccctga ctacaaactc
119221 ctcgtctggg ctcattttcc gttgggactc aggctgtct atatgtgtgc attcaggctc
119281 ttggcctggc ccaccctccc agaccggccggttgcaccacc ccatgttct cactgggt
119341 ggtgcacccct gcccggccac ccaggccctg aaccccatgc tgccaaactca ggaatcccc
119401 ttctccccc ccaccacccc atcccccagct cggccatgt gtaaacctgt gttccccc
119461 cgaggcctga ttccagaccc ctcacccctc ctggccctcc ctggccatcca gcctgccc
119521 ccctgcccata ccctgcccata ccctgcccata ccctgcccata cagcctgtgc gcaccatcag
119581 gtccacccat cccatgacaa ctccagctcc tacacttttac tcttctggc cattatccat
119641 aaaccaaggt ctcacactgg aagtgcacac tgggtccagc cttcagagtg ttcccttc
119701 atcacaagct tatgaccat gcaggggatt taaaagctt gaatttggc ccaaacttaa
119761 gaaacaggag agttcacata aacgaccaga ttccagact ctcttggaaa tcagaaaagc
119821 tggcattttcc agcaccatc gacactccat cttagctggc gctggccggg gtcctctgtc

123721 catgttaattt ttattaatta gaatcagttt cctctcccca ctcataaaat attaaagaaa
123781 ctccagaagc acactaggct gaaaagcact cagttggag cctgctggc gcctgtctgt
123841 ccaacccac agctgtccag cctggggct tacagaaagg ttgttacccc gacctggc
123901 taagtgttag tcccgcctcg ctcttgaagt ctgtggatgt ctggcaaacc cagccagg
123961 ctccccctggg ggacagagag aggatcatgg ggagggctct gtccctgacc ttcaggtgcc
124021 acccgagggg cccagagcct ggttagtaccc gggtcataa gtcaagttgc cgccattc
124081 atcctggtga gttacacaac gcctgcccctg cctttccctt ctatagaacc tgctc
124141 cagttggatt gatggtaaag tagttactag aaatggcaac tagatgtta aaatttataa
124201 agatccaaag gcttaacagc tctcccttat tactctgaat ttttaagtat tcaaccac
124261 acagcctgga agtaggggt tcacaggaaa cacaagcatg catcctgttt ccttctgc
124321 acgacccagg ctggctggg cccactgccc cttccatgg aagccacagc tgccatc
124381 acagcccccc caccgtgtt tgactggac ctggtctgc actacaaccct gtcaaccc
124441 agcataaaac ctgtcagggt tggtccaaagg ttgcagtgac aacaatcaag catagaagg
124501 acaaaggata ggaagcacac aggtccatgc acagaggac ccagacttt ccaagg
124561 agctgcaag tccctgtgc ccttcaaactc acctccctcc gaaatacaca cacagc
124621 accttgccaa gagtccccca gatgtggca tctgctaaga ttcccatattc tgaaagtatt
124681 ctaactccaa gattcttctt cccatcttc cttttgttaa tgaacctggg agatctgt
124741 gacaacgggg caatggccag gcagcctccc ctgctgtac agtgcctgg tggatgac
124801 gtctcacagg gagagctgca cagtgccctg agtcccagg ggcctccctc tcagagct
124861 ctcgggcaact ctcttgccctc tcccaactc ccaggaggt ctctaacaac cttggag
124921 ttcaaggaaag ggggctggac ccctcaactac tgaacagtcc caggagtggg aagtgt
124981 gagtgggatt tgaggcaact cgtcctcacc actggacagt cccaggagtg ggaagtatc
125041 agtgggattt gagccactcg gtccctacca ctgggggctc ctttacagtt cagcatcc
125101 gaggattca cacactgact tttccactt tccaaaagg acggcagg
125161 tccatgggt tagaaacag ctgtccctac cccatctcc accagcttcc
125221 atggatcctt cggcagggtc cagccactcc cgtccttctt ctggagagc gaccctt
125281 ctaggctct gtcttgagc tatggcttga gctggcctt ccagggttct gctgag
125341 gtgaccacaa cagagtgcc tctgagctcc acagaacctg atgactttgc
125401 taacccccc tggaccttgtt ttcaggccct acagacccctt actccagaaa
125461 tggcctaagc tggccggagc caatttccaa caggaccaag gagcccgac
125521 ccagcgaggc tctaagaaaa gctagactcc actttgaagg gcagggtgg
125581 ggcaggaga gtaagactg gacgctgagc cagggccctt ggttcaat
125641 ccagtccctgc ccccttgca gcccccaactt ccctcatctg caaactgagg
125701 cgttaacctat gaattactgc ttatccctt tattcatc cccaaaaaaa
125761 acgttggcag gtaccgatca tccttctaaa actctactag gaccaagg
125821 accaaaaacc agaaacccac tcctgtaaaa taaaggccaa ggcgtccca
125881 ggcctccat ggggtcccaag cagccctctg tccctggcgc caccgg
125941 ctctccgggc ttctatttgc acctcaaaqc ctacaccact
126001 agtgcgtaa gtagcagaga gctggtaaag atctgtcagc accaaactga
126061 gactcaacca gagagttaa aatgcaactg aaaagggaaat cattttcat
126121 aatgctttt ttcttcttt tttgagacag agtctcgctc tgcacccag
126181 atgggcacaac cctcagctca ctgcaaaactc tgcctccat ttttaagcga
126241 tcagcctccc aagtaactgg gattacttgtt acacaccacc
126301 ttcttagtag agatgggggt tcaccatgtt ggccaggctg gtctcaaact
126361 ggtgatccac ctgcctcgcc ctcacaaagt gctgggatc caggcataag
126421 cggcaggatc aatgtttt aatataaac agctttatgg tgctgtgatt
126481 acatctcagc catctaaact atatgctgaa agtgaacaaat
126541 ttttagaaca ttttcatcac cccaaaaaaa aacccgtac acattggcag
126601 tggggccccc aaccttccct ccccaaccca gcccggc
126661 ctctatgaat tggcctattc tggacatttc atatacgtgg
126721 ttgtgactg gttccctca ttttagatgt tttcaaggtt
126781 gcaatatttccattt tattgtcaaa taatacc
126841 ttatattatcc atttgcgggt ttagtggacat
126901 aatgctgcta tgaacatctg tggacaagtt
126961 aaagtggat tactgggtca tacagtaact
127021 cattctcagc gccccca
127081 cattatttctt aggcttctt
127141 gcttcggctt
127201 gtcattttttaacttttga
127261 cattttctttt tattgtac
127321 agatatacaa tttatgaata
127381 tagtgtccac tgaagtcatg
127441 tgtccacttg tgcttctgg
127501 ttactccatc attattaag agtctatag
tttttagctct tacatgtaga

127561 atccattac agtaatatt tttgtatgtt gtgaggaagg ggtccaacctt cattttcca
127621 tatgtggata tcagggtgtt ccaacaccat ttgtgaaaaa gaccattttt tcccccaactt
127681 acttgttgc gacccatgc aagaatcaat atgaggatgt atttccaaac tctcagttct
127741 attctaccaa tctatatgtc tattcttcg ccaacaccac acagtcttga ttactacagc
127801 tttgttagaa gtcttagaat tacaaggat gaggccctcc tctgttcttc tccteaatag
127861 catttgggt agtctggatc cttgtatTTT ctgtatgaat tttaggatca gcttgcaat
127921 ttctgccaca aagctagctg ggatttgata gaattaaatg cattctggag tcatgtttgg
127981 tacccatag ttcacccctc aggccaatgt ttagaaaca ttggctaca aagaactct
128041 gcattagaaa ccaagggtt ggtcagggtt ggtgctcac accctgtat cccagactt
128101 tggaggctg aggaggcggtt atcagggat caggaggtt agaccatctt ggctaacacg
128161 gtaaaaccat gtcatcttga gatgggagcc ctaagtcttat caactaaacca ggtgagccac
128221 tgitgacaaC tctggccc tggaggggtt gcacagaggg atggcttcaa ccctgtgt
128281 agagaatgcc ctctggagcg tctgtcttag ttctgttgg catccctccc cagccccccc
128341 cacccgcac tttacgcagc ctggacttga cttttaaact gagacacccg gcaagtgggg
128401 aaatcacagg tcgaaatagc ggaattgccc tctgcatcta tgctacaggg ataatgggg
128461 ttacaaatca gaagacattt agcatatatt ctggaaaaaa aaagctatgt aatatataaa
128521 ctattctata tttaacttta ataaatgggg gcaggcaaaa aattgccaaa caagcccaca
128581 acaatcagac aaagagtata ggaagaaaaa aggaattaca gtttaattgc actcgtttcc
128641 taatcttcg aatcttcagt tagcatcaac agttacttggt aagttctaa aatcaatcaga
128701 tacaagctga agaattctac tcaaaattttag aaaagtccaca actagaaaaa ctaatgtata
128761 ccttccaaat gactgaaggc agaccatata gaaaatatgg tagacttagaa aaacagccaa
128821 aaaaaaaaaa aaaaaaaaaa tccagaaaaaa cagaaaaatc aatttagatg acagaatgg
128881 gaacaaatataaacaataaaatctgaggcagt tagaatttg atgttaggtt ggcattatgtt gatcccgaaa
128941 ctgaggcagt aaaaaaaaaaaatccatggtgggtagtgggatggcggccatgttggaaaccccc
129001 aaaatacaaa attagccaga cgtggctca cgtccataat aaccccgagg cacccctgtatcc
129061 tgaggcggga gaatcacttgc ggcataatggtgggatggcggccatgttggaaaccccc
129121 ctgcacttca gcctggaga aaaaaaaaaaaatccatggtgggatggcggccatgttggaaaccccc
129181 taaacaaaca aaacaaaaca ataaatggtc agattattat tattattattt attttgagaa
129241 ctccagcctg ggagacagag agatacccttgc tctcaaaaaaa accccaaaatc
129301 cagtggctca cgtccataat cccagcactt tggaggacg aggcaatgg atcacaaggc
129361 caggagtctc agaccaggct acgcgcctgt ggtccaggtt gcttgggagg ctgaggcagg
129421 aatttagccgg gcttgggtt ggcttgggtt gcatgtatgtt gatcccgaaa
129481 agaattgtt gaaaccaggaa gactccatcttgc ttttttttggggatggcggccatgttggaaaccccc
129541 cattctggc aacagatgtaa aaaaaaaaaaaatccatggtgggatggcggccatgttggaaaccccc
129601 aaaaaaaacaa taaatggtca gtcacccagg ctgagggttgc ttttttttggggatggcggccatgttggaaaccccc
129661 gtcacccagg ccaggttcaaa qcaattctcc tgcctcagcc tcccaagtagtgggattac
129721 caccatgcct ggctaattttt tgcattttta gtagagacgg gtttgc
129781 gctggctcg aactcctgtac cttacgtaaa atcatgtata tctctgttcaaaac
129841 attacaggtg tgaaccaccg acacccat gacatattgg tgaccgaaga
129901 tgcaaaataa ttttacaaaaa cttacgtaaa atcatgtata tctctgttcaaaac
129961 ttttacaaaaa ttttacaaaaa acacccat gacatattgg tgaccgaaga
130021 ttttacaaaaa ttttacaaaaa cttacgtaaa atcatgtata tctctgttcaaaac
130081 cccagctca gcatgaccc acacccat gacatattgg tgaccgaaga
130141 acagtctaat aattaagcat ttttacaaaaa ttttacattt tgatgtatgttca
130201 tcacctctga ctgagtgcc acagggttgc ttttacattt tgatgtatgttca
130261 acagggttgc ttttacaaaaa ttttacattt tgatgtatgttca
130321 cagcacgtga taaaatcagcc ttttacattt tgatgtatgttca
130381 cagctgttag gaaacgcgtt ttttacattt tgatgtatgttca
130441 ttttacattt ttttacattt tgatgtatgttca
130501 aacagacaat tcatgtgaag ttttacattt tgatgtatgttca
130561 aaaaaaaaaa atggaaaaaa ttttacattt tgatgtatgttca
130621 gccagccagg cggccctctg ttttacattt tgatgtatgttca
130681 cctgtccaga atgtcccaca ttttacattt tgatgtatgttca
130741 tgcacacacc aatgcggggg ttttacattt tgatgtatgttca
130801 ggacacagag gtcagtggtt ttttacattt tgatgtatgttca
130861 tgacgagaac tgacactctg ttttacattt tgatgtatgttca
130921 cagcatgaaC cccacgtcg ttttacattt tgatgtatgttca
130981 agggcgctt ctactaaaga ttttacattt tgatgtatgttca
131041 caaaaaaaaa gattttctt ttttacattt tgatgtatgttca
131101 gtggccttga gacgtgtgtt ttttacattt tgatgtatgttca
131161 agatcctgc aacaggccgg ttttacattt tgatgtatgttca
131221 ccgaggcagg cggatcacct ttttacattt tgatgtatgttca
131281 accctgttctc taccaaaata ttttacattt tgatgtatgttca
131341 cagctactcg ggaggcttag ttttacattt tgatgtatgttca

131401 gagctaagat cgccaccgctg cactccagct tgggtctac agtgagagac tcttgcctca
131461aaaaaaaaaaa aaaaaaaaaaa agaccctgtc gacaggccct tccctgggct ggcacctgccc
131521cacaccagct ggaaccttaac acctttgac agctgttccc tgccagcctc cctgcctct
131581tggaggtctg cacacaccgc ccctcccttc ctccaaccag cccgggagac cagcacgcct
131641gagcaccagc actccccagag acagagccgg cacgggtgggt ctgtctgctc agggggctgg
131701ggtggcaggg gcacatgtct gggagttgcc aacatatgt tccttaccag attccccagg
131761ggacccaccc tggagagcaa ccaaaccctg caaggctgca ctccaagaac tgagggcctg
131821gggaccagg accgtcccc agcattcaca gcaggctcac ccacatccac caggagatgc
131881tgcccaggt aactgcccgc accccctccc aaaaaagggtc gggccggcg ccttccctag
131941gttgctcggg caaccatcac aggcagcac aggccggggc ttgaatagca gaacctattc
132001tctcacagtc ctggaggtca gagctctaaa attaaggtgt gggcagaggc gtttctctc
132061gaggcctctc tccttgctt gtgcacagca tccttccct gtcccatcat ggcgtcccc
132121ctgtgtgtac ttctgtctta atcgcttctt cttaggcac cagtcattca ttttgaatca
132181gtgactcat tttcaggtaa ttatctctt aaagaccctg cttccaaata cagccaatat
132241cctgatctac gggagtttag gacctaaca aaggactgg gggatctagg gaggcgtca
132301ctcagcccat aacagcctcc aaggcacaag tccacacaga atcgagcag gctgctggca
132361gggtccgggt gccagagctg ggcgggctgt tccttgggaa gccactctgg acaggccgga
132421aggtcaatgc caggatcaact ccccaggggc cagaatccca gcatggagga aagctgccc
132481tccaggaaag taggagcccg cccagggtctg gccacatggg atgcccattt cacccttct
132541tctaaaactgg gcatttcac atggggacac gggcagtga gtgcaggcgc aaggagacga
132601gaatgtcga cagccgttag gtgggtggacg ggcctcagga cccctccgccc akgcaggc
132661cactctctca cggtccacga ggagccctgc tcctagagcc acgactggag ccagctgagg
132721ctgggtgtctc acattctgc agagaggctg gctaaggctc ccgtgtgggg gcaactcccc
132781tgcccactgc accccagaat ttctggggc atcccaaggc tgagtctacg cagctgttgg
132841aagacggggg atgcaattt atagtggtgg cgtttttaa acagcatcca ataccctttc
132901tggcagct cagccggag agcaggctcc cagtcctcag gaccaggag gcaggctcg
132961ctccttccca cagcgggtgc tcctccagga gggcccaagg gctcagaggg aaaccagtgc
133021ccaccaaggg cccagctgcc tcggctggctc ccagcagtgt ggggaagaca cactaccga
133081cggacgtccc ctttagaagc aggggcagac aggaactccc atggggagag cacaggctgg
133141gacgcagccc gcccaacgccc agtcccccca tcctactgt caacaaggat cagccttcc
133201tctctcaaag cctcggttcc cccagagggg gctcaggtt agagttgatg tgagactca
133261gtggggcatg tcctccatgt ctgctgtgc aaagctgccc agatcaggca gtttccctag
133321ccagaggacc cccctttca ctccatca caggctgtca tctcccccggg tccctgggg
133381acaggccct ctcaggaccc ctccggcctc atgcacac acgactgtg cactgcctca
133441ctctgttc ctgtgaattt ggcttcttt ctgtgtgtc tctctggggc ccactgtgt
133501atccaggggca cgctgaccgc cgaccgtctt ctggccggct gcccagctgc cactccggg
133561cggcagccc tgccagggcc gaggctgtct gcttccctt ctgtccagag ttcccaaacc
133621aaaaatcagg ccataccatc tgacgaagga tcctgtctgg ttatattat ggaagagcca
133681gggctctgag gcccagggtt tctgggtctg ggccaggctg ctctgcaggc cttcagacac
133741tgcccagccc gtgtcccac ccacccaggg tggctccggg ggagggccgg gggcaggggcc
133801ccagactgg tgagccaca ctggtgactc tggatgtgt attatttcta gagccaggac
133861agcagccgccc aggcccttag gcccctctgt gggtggttgg cccactggag agccggagca
133921gcgaggccag caatgaccaaa aggcccgta tcctactgtc aggcaggctc aggcgcaccc
133981gcctccctg gatgccaact cagattcggg gagactgaac tcattgggat aagtgcuccca
134041cagatcctgg tcctcacaag ggacaaggcc accagggttc acactgggg acagtggcac
134101cccagggttat gaaaggccac accaggctta tgccatgg ccctgcctgc acagcacagc
134161acggcacggc atccataccg cagcacagcc cccctcccg gcagtccaa gcccgggatc
134221tcagggcccc caccagaaac cagacagggtc agcagggtgc agtccaggc caggtctga
134281ccttgcgtcat accaggctt ccccgagagc aggctgcagc cccctcagct gttctgact
134341ccacaggact gcccctcc tcggctggctc tgctcacacc cacactccct gctcagagac
134401ggcctgtccc accggccctgt tgaaatcac acccaaggac agcggaaacga tggaggaggg
134461actggaagg tgcttcggg ggtcaatct gggcaattt gaacacaata ataaatgata
134521gtaataggtt agaaccatt gaataaaaata aatattccaca gatcgttatt ggtgaaaca
134581aataactgaa taataaaaat gtggggagggg aaggcagct ctcttcaca ggttaattca
134641actaaaatg taaaagaaaat gatgagatag aaaatcatgg tagtaaaccc catcgccaaac
134701acacagtcg gactgcggc aggtacgact tatccgtgg ggtcaattt agtgagcaaa
134761agataacaa gaaacagaat gctcacatca cgtcacagca gctaccacg ggagatttaa
134821tcatcgcaat gggaaaaaaat agtggaaagag tggttgtct atcaagccac tcatctgaga
134881cctgaaaaaa cctcacagtc acagccctag agtcccccgg gcatccagtg gaatcccagt
134941aggttaacacc ttggccaaat gaagaaggc atcaactgaca ataggagaca actggacatc
135001ccaatgtgac agtcccagcg tggtgcccta cggaggccac agtgcctctg tgacagtct
135061gccagaaacca cgtaaactca agccagtc aaggcacacca gaaaaaccca tggggatgca
135121ccgtctacaa aacagctgac cagtcctt caaaaacatc caggtcatga aacctcagat
135181gggagggaaac gggagagaca caacaactaa atgcccacgt ggtatccctg gggagccctg

135241 aaccagacaa cggacatgag tcggaagact gacaaaattc agtcagtct ggagtgctac
135301 gtttagaatg tgactgttt gttcccacca aaactcatct tgacatttg tccccagtg
135361 ggcagtgtt gaagggtggg cctggggta ggtatctgaa tcatggagc cgatccctcg
135421 tgagtggctt gatactgttc tcaaggttagc aaggaaattc ttgctctgga gaggctttag
135481 tagtccccga gcctgggtt ttataaaagc aggacactcc ctgggtttcc cttcttgc
135541 cacatgggct ttcccttga ccctctccac cacgttatgaa cacagcacaa aaaggcctca
135601 ccagaagcca gggccatgct cttgaacttc ccagcctgca gaaccatgag ctaaataagc
135661 ttcttttctt tataaattac ccagtctcag gtattccatt atagcaacac aaaatagact
135721 aaaactaaga cactaagtac caacgcttag tatcaaccta caaagttcct ggtttcagtc
135781 cttgttctac gttgcataca gagaagctg ggtgaaggc acagaggaaa tctatgctat
135841 ttttgcacta ttcttgcataa ttttccaa acagcataga tctaaaggta tttcaaaatt
135901 aaaacttaaa acacataaaa ataaaactgg gcccagcct ccctcagtg tctctggccg
135961 gtcatttttccatccatggct gccttccctt ggagtggagg cgccactg atccccagtg
136021 tccaggagc aggtcctgta gaagggtggta gttccaggg atcggggagc aagtacagag
136081 gccctgagaa ggaagcaagc ctggcgtgtt ccaggaacac cagggagaac cgtgtagccg
136141 gagcaatgct cttgtgcaga gggcctatgc ctqagccccc ctgtgggccc tgccgttccc
136201 ttcaggggcac tctccaaaccc tctgcccgtt gcatggggcc acggatgcc caggagctct
136261 agggctttt actatgaggt tgcttcagggt tcagtgagt ggcttcatgat gcgagccacc
136321 ctcactggcc aggaggggac ttcccttcaca ggtgtaatg ggactgagcc accagcatga
136381 tgcgtacacg tgggtttga ggattccac tcacatcaga cagaatgagt gcggggctct
136441 ggatcacagc tgcccaagcc tcaagcctgat gtggcacttg tgggtccctg catcgctgt
136501 gggccaccc cccggctctg gtgttcagggt gtgggcagtg gcaagcaca tccttcactc
136561 tggcctctg tcactgtgc ccaggggcggt gccccacacc ctcaaaagggtt ccgtgtcatt
136621 tccggcagt ctggcagggtc tggctgcatac caccgttgc ggtgtttc agagggaaagc
136681 tgggtccctg caatgcaccc acccttcccg tggctcttgc tgcctccaag ccgtgtgg
136741 catcatgctc tcctcccaagg ctggctctgca gagggtatgaa atgcctgtca caactggat
136801 cccttcttc ttgtcattct gcctgctgggt ggcaggtgt ggaaggccgg agtgtgccc
136861 ccattaaacg aggttgaatt ctgggagtgca agaagtaaca gcctcaggac tgatggggaa
136921 gatgaggctg caagatgtgt ggtttttccca ttgcctataa aggtccaaac gatatttatt
136981 tatttttttta ttatttttga gacaggatct tactgtgtca cccactgcag tacagtggct
137041 caaacacacg tcactgcagc ctggacccccc tgggtcaag tgatccccc aactcagcct
137101 cccaagttagc tgggacaca gacgtgtgca acaacagct acaattaaaa agaaaaaaaa
137161 aatttttagt aatgtgggtc ttgtctatggt gccagggtgt atgtgtact cctgggctca
137221 agtaatccta ctggcttggt ctcccaaaat gctgggatta catgtgttag ccaccacacc
137281 tggcccaaacc attaaattta acacgaccag atgaggacgt atggctcaca ttttcttgc
137341 cactctggc aagatccaca ttgtgtgtca tcattaaatg agcacggtaa ccaccccttca
137401 gggacaacg gaagccggca ttcccgggggg cagagagaag cggcggagtc aggagaaatg
137461 ccacaaccag cactaaatac atggggctgg ggaatgcggg gcaagaagat atacagcatc
137521 tccacacaaa caggtccaa agagccacac attgtgagcg gtaacagctt tgctattacc
137581 ccccagctgt gggagatata aatatgaaat tagctggac agcacactcg accataaaaga
137641 acaggaagag atagctctg agccctgagca gtaacgagggt tctctgcaag cactgagcag
137701 acagttagca aatgtgttcg tgcttctctt cccctccccc taaatcttcc ttgatacatc
137761 ttgttattcc ccatgtggc tcatgtatct tgggtataat tttttatca aattgtctat
137821 atagacagga gtcctgcagg aaatatatgc ccaggcccc atgtatccta gggacagcca
137881 gccatgcctt tgggaacagc ctagaaggac tggcgggtt gggctgggt ggggtggagat
137941 cagcttgcaccc aaggtgactc tctctccctt caagatacat ctgagagaga agcccatgt
138001 aagcaggggcc aatccaagct cccacccccc cctggagtgca agaccactgc ccagacggct
138061 tcctgaggc ctggctgggg ccctccaaagg tacccttcc ccaggccctcc agccccagcc
138121 ccccaatagc ctgcctgaa ttgttctctg cttaaatgtt gggctgggt ttttgcattc
138181 caaaagagaag ccctccctt ttcttagagg ttgtcaggag ctacgtggc cggtccccctg
138241 ggcttgggac atcctctgtt gttggggggg gagggtgggaa agcatcttc ctgcagccat
138301 gacccatcactt tggtagacagc cacagttct cactcacagc actgggttaa gtgctaacct
138361 gaagccgctc ctccctgtcc acggggaaaga cagtcgtacc ccactttccct ctgctccccc
138421 aaaaataagga acgttattgg tggcttgggt tcctgacaga ctcacacgtc cttaaaaagt
138481 gctaatttc aaaaaaggagg cagccggctg tggcttccctt gggaaaggc agattaattha
138541 gctgactccc acatgagcagc cctctgttcc ttgcagggtt tgggtgggaa tggatggatca
138601 ggcttggactt gaaaactgac atctcaatgtt gaaatgtt ttcctgtca actgcagtga
138661 gtcagagccg tgcgtgtcc tagacagata gggaaatgg ccttccatg tggactggag
138721 tggccaaacg cctccggggcc acagcccagg acaagcgtgg gacatcagcc tcccattgt
138781 aggcacaaac agggcagtgc cccatagatc ccagcccaag acaagcaccg gagttggcat
138841 tccagggttag gttttttttt gttttttttt ttttctcaag atggatctt gctctgtcgc
138901 ctaggctgga gtgcagtggc atgatctggg ctcaccacaa cctccacccctc ctgggttcaa
138961 gtgattctcg tgcgtcagcc tcccgagtag ctgtgattac aggacacacgc caccacaccc
139021 agctaaatttttttttt tttttttttt tagtagaaatg gggtttcatc acattggctt

139081 ggtggctcg aactccgtac cctgtgatcc gcctgcctg gttcccaa atgctgagat
139141 tacagggtg agccaccgca cctggccaag gttttttgt tttttgttt ttttgggttta
139201 agtatttg gcatctaaa cattccactt aatgactggg aactggggcc tgcaatccca
139261 gcatttgg gaggccgagg caggaagact gcttgagccc aggagtccaa gaccagcata
139321 gtgaggcctc atctctacaa ataatttaaa aatttagctgg gtgtgggtg gcatgcctgt
139381 ggtcccagct acttgggagg ctgaggtggg agaatcacct aaaccctgaa ggtccaagct
139441 gcagtgagct aagatgctgc cactgcactc cagcctgggt aacagaaaaa gacactttca
139501 aaaagaaaaa atatgactgg gaaccgggag aaagatgtt atgaaggcac gcagcctaca
139561 gtcagagctg acagccaga aacacgttcca gccagccaat tatttttaat gaatggccctt
139621 ccttggtga ctcaatccca tgaacaggga agatgggtgc atggctagaa gctctgttcca
139681 ggctccgagg gctgccagc agaccacccc taggggtgc caggatcaac ccgaccacagg
139741 gatgctgagc accatgcaaa gtctgaccca cccacgtggc ccaaggaggc agcttgtgaa
139801 ggtccggaaat ggtggctgg agattggca cccacagact tggggagtgg aaggggccacc
139861 ccaagccctg ccccccccccc cactgcagct tcccccggcagg gctgcaccca catccgcagg
139921 gtggccctg ccccgtgtc ccagcttcca aggtggcctg ccctgtatgc cagagaaact
139981 tcctgctggc caagcgagg ggcttgctt cttagggtcag cagaaatgt caacatttc
140041 cctgaaaaat cttcaatttcc cctcattttag tctccagtag acttaggtca tgtatacact
140101 caacaggaca actcttttcc accccaaatg ctgacagccct ctgaggcagt ctaaagaagg
140161 aaggaagagct cactgtgcag actcctggca tggaggccct cctgtcgctc tcaactactg
140221 cctctaggcg aacgggctg tggggactt ggcgggttcc cactggccgc atggatcatt
140281 gtttccactt caaaaccagg cttttagcac cctactggag gctgctgggt ttgtctgtca
140341 tcctcagagc ctgagagcat ctggcaccag gcctggctga ataaatagat tctggattag
140401 aaaggaaacc tggctttgaa aagacacaga gaagaaccag gtccctgcgt cccttccttc
140461 caggctgtatg ttctggctg ggggtctct gttgggtctcc ccaccccttc accttctgccc
140521 ctgtccatac gcccagtagg cagacctgtat ggactggcctc ctctggctc tccctgggt
140581 aaggggggca ccgacaggaa gccagggtcg gggagagaga cagaaggttt tatttctccc
140641 gctgtggtgg ccccccgcct ccccccgcctg caagctccct cgccctgcctt gcagctctgt
140701 gagttgtgcc tttgtcaact ctttctgtt gatcccttat gtgtcccagg cccctgactg
140761 atccgacatg ccctgttagcc cgaccaggca gcccctactc tcatagttcca ggtgctggc
140821 ccatctctag gttttggat gggacttcca cactccagcc caatttccac cctgaagcac
140881 acgggtgagc cagttccctg caaccaccac agactcaaga cacaaggatcc caggctggc
140941 cagaacatttcc tgagcaggc agaggagccc cgtccatggc ggctggggca gtcccgcaag
141001 tgagagtccg ggtggggagg gctggggctt ctggaaaggga ttcaggagca gccaccaagg
141061 cccacccatca cacacgtctt caaagccctt tccatattcc gctgaaacca ctatggggc
141121 ctgtgtggg agcaaggagg gctggagac aggtggggctt gtggaaaggag gctgcatgaa
141181 ggagggagg ctcctccaaac ccaggacggcc ccaactgcagg agggaaatgcc ttctggctg
141241 tggccagga gactgagggt cctgctcagc cacacgacac gtgccaagggt gcagagccca
141301 ggggaagagc aaaccggagg gctgggggtg cctcgacgca tgcagtgtat gtgaagcagg
141361 agatggggg aggcttggc cgtggggaca ggagaatgac cacaggatcc acacctctgt
141421 ggtccgtggg acctcagcag acgttttggg tgatgccaaga gagccatgga ggagggggcg
141481 gcaggggctg tgcttctgag aatgctgaca atgatggac ccgggagcag ccggagctca
141541 agggacgggc aggggcaatg ccggggcctg tctgtgtcgg gaggaggggg aggtgggtga
141601 gcaaggaggt ggggtggcag ggacggggaa ggtcccaagg gggatgtcg cctgcagagg
141661 ccacgcgtca ggaatgtca agtagatct tcagcgaggc tggggccct gctcgaggc
141721 agggaaatggg ggggggggg atggggctgg accccagaaat tcagcaagga aacggggagg
141781 cgggagacag tggcacggc agggagcaag tacaccaag aggaccatgt gcatccgact
141841 cggataaggaa aggacccgca ggcagggtgg gatgggtca gtggtaagg tccccagggg
141901 cggcctggc agagggcagg agggagggc ggcggcaga gagaggctgc tgcgtctgt
141961 ggttcaggag ggtgggtcctt gggaaagacca gctgcctgccc tgctcccagg gtggacagg
142021 gacaagtcaa gcataaaagat gcaaagggtc aggccaaaa aaggacgggc cactggatc
142081 actaacacaa ccatcaatgg gtttaacattt gcagtttagag atcaaccaga ctgagatgga
142141 acaaacaagg catccgtct gtaacacatg cctgaaacat gctaataagag agacagagaa
142201 aactgcaaac ccagaaaagca ggggtgtctgt agtatcatca aaatagaact cagagcaagt
142261 actaaatgac tggaaacttcc tggatgtattt acaaggcttgc gtggaaaggcac acagtccacca
142321 agcttacaa accaaggccg gcacgtctg ccatgggtt aaacacaagt ggctcaacag
142381 cccctgtcaaa gggaaatgtc aacggccact cccagggtt gacaggctc ctagaccaca
142441 gcagaggggc acagactcat tgactcgggaa aatggggccgg attgatttctt ggatggatct
142501 gacgacttac aaccgaaagc aaatgagttat tctttacaag catggggggag catttacaaa
142561 aattgatccc ggcctccctt gcctgcctgg cctgaatccc ccaagctgt cagccgtac
142621 ctcccccccccc acctgcctgg caatgcagaa acctgagatt qctgcttctc ggcaccaaga
142681 accaagaggt gaagttggta tggggacccc ctctgtcccc actccctgtc agtactgtcc
142741 agccagaagg ccctcccccac agcctgcccc ttctcttga ctgttccctt ccccaatacg
142801 ctgaggccccc atgtcttccca ggaccctccc tcaagacccc ggcttccctc tccagggtga
142861 ctggggccct tgcataatgc ggcataatgc caactccca cccctactga tgaagtcaact

142921 ggcagagtgc ccccaacgccc ctccactata atccccaggc acggagagct cttgctctga
142981 gccgagacaa cccactgccc accaccagct ccctcagggc tctccacaag tcccttctcg
143041 cctcccccgc tgctgttcg gcctccttcct ctctaaagtc ctctctcagc ccaacaacag
143101 aaccgggtct cttcaatcct gaacaagaag actctgttc ccctccagcc acaggaccc
143161 cttccctca cagcccttct cagcaggctg gcctctgtc ctccctccca cgcaccaatg
143221 cccatggct ctgtccctcc tccccagggc cccggggacc tccacaccct cctgctcagc
143281 ctctctctgg ccggcccccag ctgctccgtt ttccacagca ccagaggcca gggtccccctc
143341 gtgccacagt gaccctgca cgcagattcc gtgctccctt cagaagccca cctgacattc
143401 atagaagctg agcaaaccacc aggagacaga ggcaatttca gaaattcccc gaggctcggg
143461 gaacccaggc tgtccccctc actcctaattc tggcaggc ctggccctgc caagcctcca
143521 tggagatgag ctcttcctgt cccaagggtg tggtgtgaag cggcccttg gaaggtgatg
143581 agtcatgag ggtggggacc cggtaatgg gattagtgc cttatcagag aggcccagg
143641 agcttgcga ccctccactg tgctgaaaca cagaagaag gcccagcta tgagcggagga
143701 agcagccccctc accagacacc aaaccctactg gcggcccgat ctacatgtg agactccaaa
143761 actgtgagaa ataaatttct gtttttctaa aacaccctactg tggatatt ttgttatggc
143821 agcctgaacg gagggacacc tgcaagccct gggatgtca gaatacacac gttgcaggag
143881 ggggtggac gggctctacg gaactgggag atgccttagaa atgctagagg gttgagggg
143941 gaggcattca caagccctt taggaaactg agatctgcgc tggtaggg tgggtgtca
144001 ctggaatggc aagatgcaga tataaatgtc gacaagaccc aatttaaacc cagagatgt
144061 aaggctgaa aacatttggaa ttatactagg aatcaacgca aagcttaata acctaaaatc
144121 caaccactcg gaagttttt caacatttct ctaataatgt ctaagttac tggaatcaaa
144181 atcataaaat tttttaaaat gatgacagt gagaattcagc acatcagaac ctgtggatga
144241 gaccaaaagcc accaccagaa gcaagttcac aaccttctta attatagacc agagtaaaa
144301 taaaatgagc cagtgcctca gtcaagggg cagaagagag agcaaccata aacctaaaga
144361 tggatata aatagcatga gtgacaagat ctcaacttctg atgtgacat atatgtaa
144421 agatggaaaa acggtagt tttatagctg taacaaatta ctgaaatgt atggcttaa
144481 aaacacaaat ttgtccaggc gcgggtggctc acgcctgtaa tcccaagcact ttggaggcc
144541 aagggtgggt aatcagcagg tcaggagatc gagaccatcc tggctaaac ggtgaaaccc
144601 catctctact aaaaatacaa aaaattagcc aggcgtggc gcggccgcct gtatcccag
144661 ctacttgggaa ggctgaggca ggagaatggt gtgaaccac gaggcggagc ttgcagttag
144721 ccaagatcac accacagcac tccagcctgg gcaacagatc aagactccaa ctcaacacaca
144781 cacacacaca cacacacaca cacacacaca cacacacaca cattattat ctgacagttc
144841 tagagatctg aagtccacaa tgggtctcac tggctaaaa tcaagggtg gggagggt
144901 catcttctt ggaagctcta aggagaatgt tttttttt tttttttt tttttttt ttttagagget
144961 acccacattc ttggcttgtt ggccccctcc tccatcttca aagccagcac cggagcatct
145021 actgccccct ctgtacttgc gaacttctgc ctccatcttca tactttatg aatccttgc
145081 atggccagct cacacctgt aatcccagcac ttggggaggc tgaggcgggc gatcacaag
145141 gtcaggagat ggagaccatc ctggctaaaca cagtgaaccc cggctctac taaaata
145201 aaaaatttagc caggcgtggt ggaggccgc tttttttt tttttttt tttttttt tttttttt
145261 aggagaatca ttggaccca ggagggtggag ttggctgtt gcttgcgtt aatcactgc
145321 actccagctt gggcgcacaga gtgagactcc gtctcaaaaa aaaaaaaaaa aaaaaaaagaa
145381 gaatccctgc cattacactg tgccatccag gataatcttcc ttatctaaa agtcaactaa
145441 ttagcaacct gacttccatc tttttttt tttttttt tttttttt tttttttt tttttttt
145501 ccagggttctc gggactgggaa catggacaaac tttttttt tttttttt tttttttt
145561 gagactatct caaaagggtt ggtttttt tttttttt tttttttt tttttttt tttttttt
145621 attattcaat gataatgacc atatgtttt tatgttcaaa agtaatttt aattgtgaa
145681 aatgttaattt attgtgaaatc atgtttttt agttgttgg aataatgttt aattgtaa
145741 aatggaaaaac taaagaaaca aaagaatagc atccctttt agcaccattt gatgtcttt
145801 caacactttt tggctggctt tttttttt tttttttt tttttttt tttttttt tttttttt
145861 gccagacccc ctttcaattc tttttttt tttttttt tttttttt tttttttt tttttttt
145921 gctttagcag acatttctca aagtgaagat aaatcccttca ataaaggagg ccctggg
145981 agaagaatga tttttttt tttttttt tttttttt tttttttt tttttttt tttttttt
146041 gatattattt tttttttt tttttttt tttttttt tttttttt tttttttt tttttttt
146101 cgcaagccag cagacagata aaccgttatg gttttttt tttttttt tttttttt tttttttt
146161 cgcggagccca gaagaagcc aataaaaatg cggagatgc cccggccac gaggccggag
146221 gatgtttttt ctttccatc acaagccca aacacaggcc tttttttt tttttttt tttttttt
146281 gcccggccccc tttttttt tttttttt tttttttt tttttttt tttttttt tttttttt
146341 ctggtccccct ctgtcaagaaac ctttccatc tttttttt tttttttt tttttttt tttttttt
146401 ctttccatc tttttttt tttttttt tttttttt tttttttt tttttttt tttttttt tttttttt
146461 ccaggcttgc tttttttt tttttttt tttttttt tttttttt tttttttt tttttttt tttttttt
146521 acttggatc tttttttt tttttttt tttttttt tttttttt tttttttt tttttttt tttttttt
146581 acctctgtt tttttttt tttttttt tttttttt tttttttt tttttttt tttttttt tttttttt
146641 gattggaaaat aagataatga aggcaaaaggc aatattacatc actttacaga gttgtcaagg
146701 cagattggaa aacaggctcc aagagatgac atgactttt caggattaga cagctaatga

146761 gtgggtggatc cgatctgct cgacacaggc tgggtcttc cccagccccg cactgggaca
146821 ggggcatttc ttgagggtcg ccctgtggcg cccagatgcc gcagggagac aaggctgtgt
146881 gcacagatct ctctccactg ggacacact gagaacttgg gggagagaag tgggtttcaa
146941 gcatctccct tgtgaaccac tgggtttaag tatgtgaac ttcatgttg acgcagaact
147001 tgagacaac ccattctctg gaaaaagaca caaatgaggg caggaccct gcaccatgt
147061 ccctgtgccc ccagcagtgt cctctcatcc agcgtgtaa gaaggagtgc agccggagag
147121 cccgtggcag cagggcgtgc gggatgcccc gtcacagtg tgctgtgcct gggagaggag
147181 cgagggcagg cccccagctg cttggaaaca gtcagctgc acacccccc cctgggtgcc
147241 ctccaaggc agccctcagg ccccaaggctc cgggactccc tgcagaggaa ctgcaagcca
147301 gtcctggccc agcacagccc ttccgcccagc ctctggccagc cgtcaggaag gggagtgttag
147361 atgagtagg aagccacaca ggacgacaag cgggacttc tgccctcac cccactgcac
147421 gatatggaaat agacgggggtt gatggccggg tgccgtgtc catgcgtta atcccagcat
147481 ttggggggc tgagggtggc ggtatctgat gtcaggagat ggagaccatc ctggctaaca
147541 cagtggaaacc cctgtctcac taaaaaaaaata taaaacatta gcccggcgtg gtggggggcg
147601 cctgttagtcc cagctgtctt gggcgcacaga gcaagaccgt ctcaaaaaaaaaaaaaaaa
147661 aaaaaaaaaaa aagcggggga gtgataaggc cccggacac ccaactgcac gcaagattct
147721 ctgcggaggg agagctgcct cttccctggg gtcctctggg ccaggaagga gcccctgagga
147781 ctgtggccc agtgggtccg gttcccaac ctccacccctt ccagcagcac cgatagatgt
147841 ggcacacaaag gtgcattcac atttgcacc tgggtgttgc tgcagaccct ggggtctatc
147901 tcccgggggg gcaaatggca gtaagcaagc aaagaggaag tcgacagaga gtgacgcac
147961 cgtgaggaca acagcacaca gtattgttga ggggggggggg ggacttcctg gaagagatga
148021 cactggaggt gaggcctaaa tgattggaaa gaggcagcga tggagagatc tggggaaagt
148081 ggctccagca cagagaacac cacatcgaca ggaccggg tcaacacaga cctgcatgt
148141 ttggggaaag ggaagaaggc caatgtgat ggtgtgttgc gagaaaggg tgggggggg
148201 cgggggacac gaggggcagag tgggtgggctt gggggggggta cagggttca gatgtgttgg
148261 cgggggcaga ctgcacagag cttccctctc tgggggggggg agattggatt ttccctctc
148321 tgcagtgggc agcccttgg gggggaaaga gacctgatcc cccctgtcgt tgctctgg
148381 atgttgtgag aattgttggg ggcaggatct ccctccctct ctctatgaaa ttccctgttca
148441 gcagggtatct gcctatggct gtatttctc tggtagaaat ccctccctga ttataaacc
148501 ctcacacat agagaggatcc ctgggggggg gttcagggtt cagtaggtat gttgagac
148561 gcctcacctg catccccccac tgcctgcagc caccctaaatg cactggattt atgagcagct
148621 agctgaggac cctgaaaaat aaataacacgc atgcagattt gggaaagaaga cttagatgt
148681 aggttcagca aaccccgaga gtttactt cattctctgg tattctctcgtt cctagataca
148741 agcaggatca aactccagaa gggggcactg ggtacagaca gacagagctc gagacaagcc
148801 ctgttgttgc agtcaagga tagaaaaaagc atttccgaat gtcagagcc aagccaaac
148861 ctcatatcc tttccttattt cttttttctt ttcttcgtgt tctcacacgc cagctcatgg
148921 cagccccacag gcacccaaaa ctctaagagg ggaccccttgc tggtagctg gaagtgtgt
148981 ggtccttagaa aggtggagtg aaacccctt gctttttctt ctttgggtcc tcccaggagt
149041 gtatgttgcag gaagtgttgc agacaggcagg gtcaataaaag ccccaagctt ctggccagag
149101 agatataaca aggagccca gggaaacttggc cagggtttagg aagatggcag agacagcaga
149161 gcttggggaa ctggcatcag aaatgttggg agaaacttcc gggctgtgc ttgagccgt
149221 cacgactggg tctgaccctt aacaacacgc catagtttctt gggactgtac tttaggagg
149281 gaccaatggc caggtccctt acgtgtccact ggggttgggtt tacatggat tgatcagaat
149341 agactgtcac aggtttttttt aatggaccctt acatgttgc accatccaca ggcacattgg
149401 aacttgcac cggaaacccaa ttaaccccaag gtttattttt ctgtcaaaac aaaaataat
149461 accttctcca tagatccaa acaagatcca aagtcttata ggttataatc aaagtgttca
149521 ggatacagtc caaaattttt tggcatacca agaaccggaa aaatctcaac tcctccagg
149581 aaatacaacc aacagataac aagactgaga tagcacatat gtttggattt actgacaat
149641 atcttaaacc agtattttttaaaaacttcc ataaatgttgg ggcaaaaactt ctggcaact
149701 atggaaagat agtctcagca agaaaacggaa gacatattttt taaaaggac caaacgg
149761 ttttagaact gaaaaataca ataacagaaaa cttaagaac tcactggata ggctcaagag
149821 cagaatggat atgacagagg aatgtgttgc tgaacttgc gatagagca tagaagtt
149881 ccaattccaa ctgcaggcgg taaaaattttt ttataatgtcc aagggctt aactggact
149941 gaaaataccg aaagatctt ctaacacttca tggcccttggaa ggagagaga aagaatgtat
150001 tactaaacaa atattttttttaatgttgc tggcccttgc tcaatttgc tggaaatgtcat
150061 aacccatcca attcaacaga tcggccaaacc tcaatgttgc gaaatccaag gaaacacaat
150121 cccacacgtt ataaacacac tctgtttttt atagacaaaaaaaatgttgc aagcag
150181 aaaaaaaaca gtgcattact tataggggaa aatgttgc gataactgc gatgttgc
150241 caaaactgttgc aagtccaaag gaagtgcggaa acattgttgc aatgttgc
150301 agaactgtca acccagaatt ctatctccag caaaataca cttcaagaat taaggag
150361 taaagacatt ctcagatgaa gcaaaaacttta tagaatgttgc gggccaggaa
150421 gatgggttgc gaaaggaaat cttcagacag gggggaaatg atgttgc
150481 acatcaggaa taaaggaa gcaacagaaa tacctgaaaa aatataacag attaggct
150541 tgcgtatccaa aaagggaaat gataatgttgc tctgtatgttgc tttccatgt atgttgc

150601 ataaaagaca acaacattca ggggaggaca gagggaccta tatgtatggca gggctcaac
150661 attcaactt aagtggtaaa atattgattt gaaaaagact ctgaagagtt acataggtt
150721 attataatcc ctggagcaac cagtaaacac tatacaaaga tatacaaacc aacacaatag
150781 gtaaattaaa attgaatacc aaaacatgct caaatagaaa aagcaaaata agcccaaagc
150841 aagtagaagg gaggaatga taagagtaaa aatcaatgaa actgaaaata tgaaaataat
150901 agagaagatc aatgaaacca aatttggttc tttgtaaaga tcagtcaact gataaatctc
150961 tagtgagact gaagaagaga aacagaaaaa aatatgaattt accagtatca gaatatttt
151021 aaaaaggata tcactaatga catcacagcc ataaaaaaga taatggaata gtataaaca
151081 ttctatgcat ataaatccat cagtggtgaa gaaatgcattt atttcctcaa aaactataag
151141 ctacccaaac ttacccaaaga caaagtagat agcataaaaaa gttctgtac tattaaagaa
151201 aaagaattca tagttaaaaa tcctgcaaaa aagaatttc caggtctaga cggttttact
151261 gggaaattct actaacctt cagggaaaaaa tggcaccaat tctcccatc attttatgag
151321 gtcacacat ccctgtataaaaccacaa gaaagacatggtt caataaaaaa aaaaacccta
151381 caaccaataa tccctcatga acctataaac caatccctt caatgaacat agacacaaaa
151441 atcctcagca aaacactggc aagtctgtc cagccatgt taaaacagcag agtgcacac
151501 agacaagtca tggttgcct ggatttgcag agtctggcac aatatccaaa agtcagtgt
151561 acccagcata ttaacagact gaagaaggaa aaaaaatatg cttatataa gtgatagaga
151621 aaaccatatg atatgattca ttcatgacta aaaattaaaaa actctcaact aacagagaaaa
151681 gaaacgtgct caacctggta agaataatcta caaaaaaaaaact aaagctaaca caataaccga
151741 tcctgagaaa ctgtttccc ctaaaataag gagcattttc acgactccta ctcaacccct
151801 ggtttagtcc tagacagcac ctctgtacac acccagcacc cagcacagac cctcaggcca
151861 gggcagggaa gggccccaga gagttccccca aaaaaccctg ttcctgggaa aggtggatgt
151921 ccaggaaggg tgactggccc ctaaaggctt ggggaagccg gggccctctt gctccagcc
151981 catgaagctt tgatctccat tgccaagcccc ggtggcagca gagggcagg cagatgaact
152041 caaccaggaa aaggacaccc aggaaggaa caactctatc agatcgatc tcatagcacc
152101 aagtgtccc acaggctctg aggacatggc tcttatttca ggaccaggcc tgacaggagc
152161 acagcctgag ggatgagagg gcaggccgtt cttctgtgg gtggggatt ggccaccagg
152221 gccagtcctt catctgtgtt gggcttggca ggatcaaccc tggccttag tccctggagg
152281 tgaggccagc acgttagtaga gggcagccg agtaagaaga atctctcggg gaagtcaatg
152341 gccagccagc ctaagcccg cccagaccctt cggccctgcta gcatgagaca gacacagggg
152401 gcccgcactg cagagctgcc cccggccac cggccctgtcc ctgtctgtctt cggccatctg
152461 ccccagcagt gctgtgagag catcgacacc ctcggccatc agccttcagg gacggcaaga
152521 cctggggccc acaggaggcc caccggccctt ctcggccatc gagcaccggg ccaaccctgg
152581 gttggccactg ctcagactc caggacactt ccatgacaca gaagcagctc tgcccatgag
152641 ctccaaacag cccggcccg gggcaccac caccggggg cttccgggtt catgtgcac
152701 catccaaatc tagcaccatc acgtatgcca ggtggggac ccggccgtcc caagccctgt
152761 acctattctt ctatgtccct cacggctgtcc ccacagaatc tctgaaacac gactcagacc
152821 tgcactgccc tgccaaatgc cttccagccg cttcccaaccatc catgactgccc ccggccctt
152881 ctgatttctt cttcactcaa gtggccagccg tggagatctc ctctgtgtcc ttggcacact
152941 tcaactcagc acctgcctca gggcttttc acctgtctt gcctccaccc ggaatggct
153001 tcttccagaa gtccaggtcc tagctccat gcttcccttcc aaggacctt gtaagggtt
153061 tcccttctcc atcacactcc acactgccc tctgtccat tttttgtaa caccatttt
153121 caatcatcga cacgtgggtt ctttattttttt tcaactctctt ccccttgcacc agggatctt
153181 cgagagcaaa gttgtcttgc ctgggtccac gtcaccac cagtgcctgg catctcacag
153241 gcatggggat ggtacggga ggtcgccggc cttcccgagg cttcccttcc ctacttcggg
153301 atgggttaag gtttcttgc caagggttttgc gttggcaga gtggccaaatg tttggccgtt
153361 tctggggcga caacaaggaa aaaaacagggg ttttttttttttggatggggatggggatgg
153421 tccccagtttta gtaaaggcatt caggctgtcc ggtggggatgg ggaagacagg cgctgagat
153481 tgtctgacac aattgtcttgc gtcaccaccc tggggatcc gttggggatcc gtcaccagg
153541 ccatctgaca aatgaccggc ttggggatcc gttggggatcc gtcaccagg
153601 tcgaaatgcc agctggaaatg gtcacgcaccc ccaagctggaa ggggagccgc gcaaaaaat
153661 ggcaggctga ggacagaac gggcagccgc cactgttccac gagcaagggtt gcaaggcc
153721 cagagccac aggacgttcc gcaatgcgtt gaccatgcctt gcaaggccccc ataataagaca
153781 gcttcatt
153841 acagtttctt ccaacatgtt gcccatgtt atgtggattt gggaaattt ccaaggcc
153901 tcagagaatgtt gcttcatt
153961 agtggaaatgtt gggcgccggc tgccatggaa ccccccacccctt ctgcagggtt tggccctt
154021 tgccctcttcc tccttccatcc ccccgccggc ctggggatcc gttggggatcc tggggagcc
154081 gacagccact gaagccctcg gggactgtt ccctgcctttt agcatcccccc tcagctgccc
154141 attcatttcag ccatttttttcc agggaggacag gtcaggccat ggtggggatcc tgggtccagg
154201 gaaagatacg aagagggtt gggccggcgtt tccctggccctt gggatccctt cgtgcctt
154261 gcaagccctgtt gcttccttgc tggactgg agttctgtt gtcaggccatc cacagggaa
154321 tacatggaga gatggggccat ctcagccatc agggcggaaac accaatttgc ttggggatcc
154381 tgggttccccc gggtagggc cagtttaggg ccagcatcc tcccaagggt tggcagg
154381 tgggttccccc gggtagggc cagtttaggg ccagcatcc tcccaagggt tggcagg

158281 aatggacaac atcgaaaccc gatacatcaa gaaacagcaa tagaatagca tttcaggact
158341 tcatggcaac cccagaacca gccagcagcc taaagccta ccgggctcc ctccctcagga
158401 gtggcctgtg aagagggac gcacacgc tCACACTCAT gcacacacac acaccgagac
158461 agaggaggc ctggccagag caggggtgtc agcagaggg acgcacacgc attcacactc
158521 atgcacacac tcacaccag acagaggagg gcctggccag agcaggggtg tcagcagagg
158581 ggacgcacac gcattcacac tcatgcacac acacactcac atcgagacag aggagggcct
158641 ggccagagca ggggtgtcag cagagggac gcacacgc tCACACTCAT gcacacacac
158701 actcacaccg agacagagga gggctggcc ggagcaggg tgccgcaga ggggacgcac
158761 acccattcac actcgtcacc acacacactc acaccgagac agaggaggc ctggccggag
158821 caaggggtgtc agcagaggg acgcacacgc attcacactc atgcacacac acactcacac
158881 cgagacagag gaggcctgg cggagcagg ggtgtcagca gagggacgc acacgcattc
158941 acactcgtgc acacacacac tcacaccag acagaggagg gcctggccgg agcaggggtg
159001 tcagcagagg ggacgcacac gcattcacac tcatgcacac acacactcac acccagacag
159061 agggggcct ggccggagca ggggtgtcag ctccaggacc aggatccc aggcagaacc
159121 tcatggtggg gccccgggtt cctcacacac cccacacccc tgccatgcc cagacctcca
159181 gcttcgggct ctgcattcac acggaggctt cggattggc gcagcctggc agcacacaga
159241 cggccaggg ccagggccag gactggcatt tccgggctc ctggccat cctggccctc
159301 ccacccgggtc ggccctgtcc ttgtcttgct cagctcgggg gaagcaccc tcgcagcca
159361 tctggggctt ggctggcct cagcctcctg gtcgcacgt ctcaactggc gctggccggc
159421 cggcagctct gagctaagtg agcgcagagg agagtgcagg gagcggcaag gccaggctcc
159481 tccccacccc gtggcgtggc gtggcgtggc cggataactc agttttccct gctttttgag
159541 acaagagtag taaatgcacc gtgtgtact cagagcccg gaagcagccc tcccacccac
159601 agcgctcctg cctctgcage ccccccggcc gttgtctct ctcacggcc cacatcctca
159661 ctggactgtc tccctttac agatgagggtg agacagcatg cagaggccca tggcttgc
159721 ccaaggggtg acggggcaga attacaagcc cctgattctg gggcttcca cgcttcca
159781 ccaagccctt gtggccttcc tgcacccagg tggggcttcc gtcggccggc agtggcccca
159841 acctctactc cagcagaccc cccacccatc agaggagttt tggcacacaca ctttttttg
159901 ccaggcccaa ccaagagctg caagggacat aggaatgtcc cttagcatgg cccgggagcc
159961 tcaagcccg ggcctaccta gaagtaactc tcttccttga cattaagcga ctcccccctcc
160021 agagcctaatt cacagccagc tacggccctt tctcccaattt attcaagcct ggaaggaacc
160081 tgaagtggg actggattcc tacacccacc ccccccacaa tccccaaggc cccacccccc
160141 cctcctcctc cccctggccc tgcacccggc aggcctcaact cactctctt cttctgttcc
160201 tctgggtcc ctgcctcaga ccctcccaacc taacacagcc tctggccacc cccagagggc
160261 tctccccaag acacgcgtcc acgtcgccat ggacaagattt aaataccctg gtttctggaa
160321 agcaactgagg ccacagggtt cttctggaaa gcaactggc cacatggc ttctggaaag
160381 cactgaggcc acagttggct tctggaaagc actgaggccca cagtgcttt ggaaggtcat
160441 gtcgggatcc cagggctgac tgcactcctg ccagtgaatg ggggtgtcc ccaaaggggg
160501 ttgttttccc accattggg gccaggatcc caggtcgggaa aacagacattt ggggtgggagg
160561 gaagccatct ctgggactt tcctccatc tccctgtcc cgctcaggg gagaagccag
160621 agctagaagc tggcatgagc cctcttagca ccaccagcc cagttccac ctccggggac
160681 acgagctggg atctgagtgg cgcaggact atgttcttaga gaggccgtac ccaggaccgc
160741 ccatggcacc cagagaaaa ctgaaaacaac gatccttcca gagacaaagc acagcccaag
160801 atggccacag gaactggccg ccccccaggag ctgtcgagca caaacgcac tggagagagg
160861 gcttgggtgc tgcacccagg agcagaacac aggaactgaa cagccacagg caccgcagaa
160921 ctgggggatc tggacagac tctcgctgt tgctaaagaaa ggacaaagag acagaggcga
160981 gggagaaaaag ctgcagggc tggcaagtc ccagaagaga gtcggggagg tgagagaagg
161041 aatatttcaa aaaacaaaat tacatgaaaat attccagatc ggatgggttc ctagaggcc
161101 tatcaataca tatataagaa aaagctccat agctagaccc acgataatga aacctaagaa
161161 catcaaagac aaagagaag ttccaaaagct cccagaggaa aagacagat cagatcagt
161221 taaacagggg ggcagcaggc agacatccaa gtcccagact gggcgtggg caccggccag
161281 cagacggggag acaagagccc aggaagtccct gctggcaag aattcaaagg ctcttgcct
161341 gagggaaagag gtcggcccat tccggggccca gagccaccc agcagaggac agctgtgtcc
161401 ctcagagggg tggcaggga cagaggccac accagagccag cagagatcta gcaacccaga
161461 agggtgacac cctggggatg gtcggagacca gcaaggccct caggaggcagg gctggggagt
161521 cggccacacgc tcagggggtc attgtcatgtt ggcattcccg gtcctgcctt aggcccccta
161581 gggtccccca gccccacaga acaccctccc cactccagg gatggcattt gagacccccc
161641 tgcgtgtcc ctccagaaac ctgtcgactt caagagaaca gaggctgccc tcttcaaccc
161701 cagtccagca aggcccaggc cagaggccac acggagaact gctgtcgac atgaagtgg
161761 gacgggacgc agtgggttac acctataatt ccagcacttt gggaggccaa gaaccgctt
161821 agtccagaag ttcaagatca gctggggcaa catagtgaga ccccatctct aaaaaaaaa
161881 aatttaaaaa attagccagg catgggtgtc cacacccgtt ctttcagctt ctttggggc
161941 tgaggcagga ggtcccatg aaccaggagg tcaagggttc agggagctac gattgcacca
162001 ctgccttcta gccccggaga gagaagacc ctgtctaaa aaaaggaaat atatttaatt
162061 tttaaaaaag aatacacatg ggtgagaaag gatggaaaga agagaggaaa ggaaggaaaag

162121 gataggaaaa gaacaaaaga agaaaaagaaa gaaattccag ctaattctca gcaacagaca
162181 attacaggaa cacaaggaca aagcccagag ccccaacccc acgtggcctg ggtccccaca
162241 gcctccagcc ctcctgtcac tctgcccagt ggcagcaggg ccacctctga gcatccttt
162301 cctcagactt ttggcgctt tcctcccaagt gtttcttg tagttctca taaatccatg
162361 ctaggtctg gctacagcta tccttgcctt tgaagatgga ctcagctgc agagtcgctg
162421 atatggtaac acaatatcac acatctcata caagcctgtc cggacgctgc ccttctgacc
162481 ccgattcctg tcaccaacac aacaggaaag gagcttgaag acccaggagg ccacagcaca
162541 gtccctcggt gaggagaga gatgcgagca ggtggcggg ggggtcactc atgagacttc
162601 tataaggcacg gacccagagg ccaccccacg gacactgtgt ggggactcct gcacctcttc
162661 ctgctgccc gaccatgggt ggcagcagc actgcagtgt tggtgtatc tatgtcccc
162721 ggtatggccag cgggcacaca gcatatgtgt gctgaggggac acatgccccca cacaacgc
162781 cccacatgtg catgcatac atcaccacaca cacacatgtg tacatttcacg cacgcgtcta
162841 ggggaggtt gctgcgcacg accagctgag actccctggc cagactcata cccagttgg
162901 gagagtggag cagacgtaca ggatggaggg tagaggtggg gctggtccac agtgagggg
162961 ggaggcaggt gttagactggg ccaggaaag gttctggaaa cccagactga aggttccctg
163021 ggtccctgg ggctcatcca ggctggggt ggagccagga catgaggaac gggcagtqat
163081 gaatctgaag ccctgttgtt cagcagagcc tccacctca cgtccctcccc aaaggtcagg
163141 caggcttac cctggccacg tcttaccct ggccattagc aaagaggccc ccctggacc
163201 aggaggaaca cagcagccca gcccctctgc cttcccatc caaggctcca gccactctca
163261 tggagactct gagctgagac actcccaagca gaaagggtca ttatcttag agcttaacta
163321 gccccttcca attagtggg aataaaatgg cagcattac tctgaagaac tgcgtatgcga
163381 gagttaaat taaatgcgtt ggcgttgagc ctcagcacac aggaggctga gcccattta
163441 aaaaggcatg agactaaagc aagacaggga ggagaggcgg ggcagcttag acaggggatg
163501 gggagagaag gagatggcgg gccccccggg gacagccagg gacggccaca gctggcctgc
163561 caggaggggac agggccact gcaaaatggg acagctgtt ctggccctc cacaagaagg
163621 ctgggggaa agaagccaag cacacgcaga ggggtctccc gcacaccct cgcctccgtc
163681 ttttggaat tttagacgtt agaagacgac aaggccact gtccagttgc tggcatttt
163741 ccccatgaga ccacagagag gcaagggaca gtccttctc agaccaggc aggtggggca
163801 atgccaaggc cccctcggtc caggatggg ctgagccccc ccaggatgac tccccctggg
163861 gaaacactga tgcgttgtt tgacgttga gacgttgggg cctggctct gagaggcggg
163921 gccccgaagg gtcttccagg aagtcagac cccagccaa aagcagcagg cctggccac
163981 aaagagagct gttccctggg ccagaaagcc ctttccacc atgactctg ggcaaggggc
164041 gtgtgtgcca tttagaacac agtgcacgc tatttcttc ctgtcttctt ccagctgaag
164101 attcacaca gcccctttt gactgggtt tcccaacago tatttctca acacacaggc
164161 ccccacac atgcattgtc actcatgac tcccacatgc agcatgcat caccacata
164221 caaatgtcca tacgcatcac atgcaccacg tgcactca cacgcaggcg tcacacaggc
164281 accatgcaga cgcatgcaca cacacatcac tcatatgtgc atctctgtg acacatgcat
164341 cacatacaca tgtccacaca tgtgcactac acacaagttt acacatgcac cagaaaaatg
164401 ccccatgcat gtgcattggca gacattaccc acacatacac atgtcatgc gtcatacatg
164461 tttatcatat gcaggatatac tatgtacagg catgcagtgc acacatgcat gcatcaagca
164521 catgcattat acatgcacac gcccgcacat gtgcacacat gcatgacact catgcactca
164581 cacatgcctc acatggcaca acatgcatac ctgacgtgt cacacatcatc ccacacacat
164641 gcaactactc acaccctct gtcgttccca gcctcaactgc ctgtttgtt actttccctg
164701 ttctgcccac tacgtggaa ggttattt tggcctttt cttattttt gtctctacca
164761 cctcatctt cccctccca gcccctccat ccacactaa ccagtagaat agaaggctct
164821 cccctctgtt gaggaaaaaa gtcgtgaccc tcttttacc cagaaaaccac aggaaggcat
164881 cgtgcagact ggactcagcc ttaccagcc cccgcagtc ttacgcctgg cttccctgg
164941 cccaaaccca ggtccccact ggtcagcacg ttacacacca gcctgccaca tctccggcct
165001 gggccacacaa atgcctggtc ccagactctt aatgttagaa gactgcccctg cctgatctca
165061 aggcttgcctt ctctgcattt cggccactcc tgctgcccag cctgacggct cccggagcag
165121 ccggtgatca gtcaacctcc ccagccttct gtttgcctt cagagccagc cgactccctg
165181 gagagctgtt gaggctgtc agcagctca ctgtccctgc ctcgttgcgtt gagggaagg
165241 agggccacggaa gggaaatccc tcccgcttca tgcgtgcctc gtgtccctt ctctctgtct
165301 gtgtctctgg gcagcagctc cccctccca gcttccagg tcccaatcc tggcccccagc
165361 cacttcctt ggaaccccaa gcctgcctgg agtccaccc cttccatttttttt tttgggttt
165421 ttttgagacg gactctgtc ctgtactca ggctggatc cagttgcacg atcttgactc
165481 actgcgagct ccacccctt ggttcacacc attttctgc ctcagccctt tgtagtagctg
165541 agactacagg tgtccacccac tacgcctggc taatttttttttttttt ttttttttt
165601 gtagagacga ggtttcacca ttttagccag gatggctca atctctgc cttgtgatcc
165661 gtccacctca gcctcccaa gtcgtggat tacaggcatg agccactgtt tctggccac
165721 ccttctctt ctgatagcac agccctgtg tcactcagga accagcttc atccctgcgt
165781 caccagtgaa caaggaggc acatgacca gcttaggca ggtagaatac gttgtcatct
165841 gactacatga acatagctca gggatgagga ctggccacaa ggtattccat gaagttaaat
165901 ccaggaacta ttgtgttgcataat gaaaagctgtt ttcccccggg ttgtgtccct

165961 gtgaggaggt aatgcacgac ctgcctgcat ccattttct gcctcagaga agagcaggcc
166021 ttagagtcaa atcgacatag aggagagcaa agccaagggt tgggatgtg ggaaaaaaaa
166081 caactgaaaa ttatttggg caccctggatc cagccacacc tgaagctaga tagttctaga
166141 tattttctt tttgcagatt aaccaataag ttttttttc tttttctga agccacttt
166201 acttggattt ccaacacttg aggaataaca cctgactaat acactctctc ttctatagct
166261 tcaaactctc catctcccg atccctatcc tcggtccctg agcaaactca agccaccta
166321 aatcctctgt gctcgaggca gggccacaca cacatctaca caagcgcaca cgtaaacact
166381 gcactcacat ctaagactgc acaccctccg cagctgatgt ctgccccacc acagcatgtg
166441 tgccatgcgc accagttcc ctccactccc ttgcctctca ctggccctc acctggtcag
166501 tctagctcc accaactccc cagcgaccct gcaaagcccc ctcttggctg ctggccctga
166561 tggccctctgc catcccaatt ccctgaaccc cctcctcttg aacctctctc atccttttagc
166621 agtgtggcag ccccgctcct aagacccac tgccctctcc acttgtctgc acatggtggg
166681 gggccaaaggc gcacttctt cccttctgc ctccctctt ctttgggcc cgccccaccca
166741 ccccaggatg tggaatccac agccctccat acagagcgt cccaaacccga attctaccca
166801 gacccctgccc ctgcacccat gaccgtcatt tccagctgcc agactcttc caagagacag
166861 ccccagtgcac aaaccctctc tccaaagctac ctttcctctc taccacccat ctgtgcaagt
166921 caaaagctgg ggagtgcctt tcctcagcct ctccctcagc ccccaggccc acagacccaa
166981 gtgcctggag atcccgccct ctccatgttt cctgagtaga ttatgtctcc tctattgaaa
167041 gatgttcccg cagaacactg gatgatccag aaatatttcc attatgtgtg gcatggcag
167101 aggcagatca caaagggaaag gtacaatacg accctttgt cattaagtat gtttacacat
167161 atgtgcattgg acgtgggaga cacacaaaaa tggtaaagc taatggcact gcagggattt
167221 ccattttctc cttttggctt atttctgcatt tctgttttat atgaaggcac atacgacagt
167281 ttttaactta aagacaggac aaaaaggatgc cagccatcgat ggtccacccc aggactgccc
167341 gtgcctggaa tgccaaatc tgatggacc cagcgcagg gggagagga gcagaggatg
167401 caccacggc ccgtgccttc tttgaatcag tcctcacggg gccatctcaa ggtggagact
167461 ctgaaaagtgc agagaagtga aataaaccgc ccatggtac acaggaggga gtgcagggtc
167521 cagttctccc tccccccaccc aggtccgcac tacaggctct gatccacgcc ctcccttgag
167581 gacaaacccct gagccacgct gggctctccc gcaggccctg aaatagagct tcattgcccac
167641 atcagacacc gggagacgcg aggaggccgt gagccaattc cagatgactt gtcaaggctg
167701 cgccagccac agagtcaagg aacgcctctgg ctgcctctgt aataaaagaa ccaggactga
167761 agggccctct gggagggcac tagacgaggg cctcggcagc tccgcaggca ggtggggccc
167821 agggccaggcg ggacagaaac ccaaattgtgg cgatgcaac tggacagaaa gggccctctc
167881 tgggctgcgg ctccctcaat cggggcctga cggtaccagg gggaggtaat ctctcacaca
167941 caagctggc ccagaccctt aactgcagggt gtccttaagg agaggtgcct ggcaagtct
168001 ctccctgggg tggacgcctt cctcctgagc acgggtgtgg agcaggccg tgcccagcat
168061 ggggtgcgtga ggactggatc ctgcctgcagc acaggactgg cggaggccg aacttgcgt
168121 gcttaggact gaggccactc cagagccaag actgacagtc tccacaaaac caaggagccc
168181 tcacctgtct catggctggc aggggctgca caggaagtga cagaggctg gtgcctcga
168241 gaggcaggca ccgagccagc cagtgtcagg gagcaggctt gcacatctga tggaaaatcc
168301 atccgtggtc ccccttctga caggcctttt gctgccttcc cttaataata gctgcgaagg
168361 gaaatggccct ggcagaagga aagacaatgg ggccccaggg gcagcagtcc aggaggccct
168421 gggccccccca ccctccagct gttaggactcc caagctaagg acagagagaa tactttgggg
168481 aggagaaaaga agcagacacc aggttcgtg gcagggtgtca gcatggctga gggcaggccag
168541 aacagccgccc agcttgcattt ggacacccctgg tcctgcacca ggcacacga cttgcccagg
168601 gccgttgggg cctgtcactg ccctgtgggg gtccttggg gagttttaa cctgggctcc
168661 ataacttggg aaaagagaac acttcaattt catctttact aactggaaac taaaagctgc
168721 cattcccttc aactctggag agaggcagca accccagat taacagcacc tgcccccgtc
168781 agggcagaaaa tcacaggtttt cggcaactgt cagcaccgt gctgcagacc agaccaatg
168841 gtgtttacac tcagcacaac ttaaattata gaggctgtaa gatgcactcc taggtctt
168901 tatttaatgc atcaattaaa agcacacaaa aaagttttac atgaatgttc ctggcagcat
168961 tattcacaag agctaaaggg tggaaacaac ccaatccatc aaccatgag tggatttaaa
169021 aaactgtggt acatccatgc aatggaaatca tatctttcca tccaaagtaa tgaagctctg
169081 atgcacatgtc caacccatgt gaaaccttggaa aacattatgc tgagtggaaag aagccagaca
169141 caaaggccaa atattgtgtt attcattttt actaatgaaa tgcccagaac aggccaaatct
169201 atacaggcag aaggtaggtc agtgggtgcc taggcctggg aggacgggg aattgggggt
169261 tactaaaggg tctgggggtt ctttcagag tcttgcacat gatctaaaat tgggtgtatg
169321 gttgcacccat tctgtgaaca tactaaaggcc attcatttggaa tttagtgcattaa taaatgggt
169381 aaatacatag tatgtgaatt atgtctcaat aacatttata tcaaaaaatgt gtaaggata
169441 ataatgatta aaaattttaa aagcatgtgt atgttgatag ctttatccccat tctttttt
169501 ttttttttga gacagagtct cactctatacg ccccaactgg agtgcagttt tqcgatctca
169561 gctcaactgca acctccaccc ccagggttca aacaatttcc gtgcctcagc ctcctgaata
169621 gttggaaacta cagatccccg ccaccacgc cggctagttc cttgtatccc agtagaggca
169681 gggtttcacc atgttggccca ggggtggctc aaactctga gctcaggcca tcctcctgccc
169741 ttggccctccc aaagtgtcgg gattataggc gtgaggcacc atgcctggcc gatagctt

169801 tttcaatca gatttatttcc tttgttat tttt gtagtgtaca ctaaacacat tattttgatc
169861 aggggtccat gggcctcacc aggcgttcag gagggttcac gacatggaa agctgtgta
169921 actaccaggaa cagaagacag gaagggttgc tggaaagaggg gtatgagaa gttggccagg
169981 cctatgggt tcccagcatt caggcaaaag cccagaggc ctcaggacc cccccactgt
170041 ctgtctgtct aatgcaagc ccatgaggcc caaggagtc cagcttcttc agccacccga
170101 tcacaccccg acacacacta gcccctgtt cttcaagtc ctttaggtcat tttcagattc
170161 aacttcagcc aagacaaaatg gtccctctt gcacctgcca caggcactgt ttcaacacgt
170221 ggcagagttt acagttgtat ggtgtgtgt atcttcgga gctctgtcc caggcccaca
170281 ggatgaacgg gaagggtggac atgggtggta gagcaagcag gaaatgtat tggtcaggag
170341 gggccttcaa tgctgggtt gggggcctgt gctttaccc caaatggag gagccatggg
170401 gacatttcg ggtggggcac ggaaggggaga ccgcgtgtt atcatggca gctggaaagg
170461 agtggggagaa acggggcag ggagaagaca gggcagatgt cctggagaca gaaggcaatg
170521 gggaccata gtcctcttc acacatcagc ctggcttcc tctggggccc atcctctccc
170581 cgaccactc ccagaggc acagggacaga cggcaggcc ggagccacag cagtcctgtg
170641 acgaggcctg gccagccaa gtcctgctcc gctgtcttg ccatagtat ttgcctgaag
170701 catgtgaccc aacctgagcc agtggaaact caccactggg ctctcccaga aaccctcggg
170761 gaacctgggc ctctcttcg cggcggctg gtgcactgcc caggcaggac ccacggggag
170821 ggcggcttagt ctgaaaccgt gggggaggcc tggctgccc ccctgagccc cgcatcccc
170881 ctttggccccc agccgttctt gccccgggtt tctgtcaattt gcaatgtat ggtgtggac
170941 ctggggatgg cctgtggag ggacaggggc acaggccaca gtcacagg tttggatggg
171001 aggacaggag cgaggccagc cgcctctggg ccacaccctc cggattcact cccaaaggag
171061 aacggccagg cggccaaatc ctcccagctg aggggtgcgg agaaaggat cagagcgctc
171121 ccccaacaaa ggctgccttcc cgccttcgca catccctcg ccacactgac ctccatccc
171181 ctcttgggg ctcggccccc acgagactt tctccattt cagactggaa atcttttccg
171241 caccttgcattt aatttttgc aggtggaaag gattttgtt gaaatggcg tttgtgtg
171301 cgtgtgtgt tttgtgtgt tttgtgtgtt cacccttccg cggcagctg ccccaggaat
171361 gctgtattgg cggggaaagc tagaggacc ttggcctcag ccatttcctg aggggctctc
171421 ggtgtatgtca gccccttcg gccttcac gtaaaatctt gtttttattt tgaccaggaa
171481 attccatitc agaacatgaa cggcctccac tggcaacaca tcccaaggcc gtcgtccctg
171541 gggcaggag agcatctac cccggctgt gcccccccc cggccccc ccccaactccc
171601 atgcccggcc aagcccccac cacagagctc taaaaggaa caaaatctgc acctggctcc
171661 gttgttccaa ggccttcac actgtggtcc ctccatcgag cccatcgagg aaaatatgtat
171721 ctgttagttaa caggcccttc gctgtctgtt gtcggccctc gtcattttt atcctcatca
171781 taaccatcca gatgttttgc cgtcatcage ttttacccat cagatgag aaacaggagg
171841 ctcgtcgcc cagggaggaa gggagcagaa aaccaagccc acgtgttggc gcgagatcgc
171901 gtcgtactcc agcgcacggg accttgcagg ggctgtatgg caaaaactgaa agtcatggg
171961 aggagaggaa gagggaggagg gagagagagg ggaggaagaa atgaccgttca aaaaaaaaaaaaa
172021 gcagcactca tctcaaaaata agagaaactc tcaggtgcct acacaggag tgaatqcg
172081 tgaagccgc caggtaatg ataggagca gcagagactc ccaaggcacc tggactgtg
172141 ggacccttctt aggggttcc atccttcgtc agtcttagcc ctccagcaca tagggccggc
172201 agatttcctg gttctcaaga gaaactggaa atcttgcgtt ttgtgagaaa ttttcaaaaa
172261 ctggaaaaggc caaacaaaac tcattgttgc gtggacatgg cccgtggccc accagtgtgc
172321 aaactccgcctt ctcacatg ggaaccctga agtgcagaaa attatctcat tttttttttt
172381 aaggcaggaa ctcaaaacac ggcttcctcg tccacatca tagatgtcc aggttgtacc
172441 accaatcta ctcggccaggcc ttccctgtt gatgtcttcac tgaaccat ccccccgtata
172501 aactcaggcc atcttccaa gaccgcagg ttcctgcagg ctcgaaaca accctgtgca
172561 tggccgggtgc actgtcccccc aaccacaggg gacaccaatc agaccacgag ggactgag
172621 accctccctt actgaaaaggg aaggtaacc cccgcactc atccccataa aaccacaaca
172681 tgggcacatc ctgcgttgcg agtctactt gatgcagag agcaggaaa ggaaagccgt
172741 gccaccgtgg ggaaatgggc ctgggagatc cattcaggcc aaggtttgc cactttggg
172801 ctaccaccac cagcatttccac agaagaagca gcacctgtt ctcaattccc ttggaaagg
172861 ggcttagatgc tctgaaggca tgggtgttagc tcccacacaa aattactaaa gacataaaa
172921 gaaaacatgt tgaaaaagaa gagacaat tatgaactgc agaaagtaac tgcacacgta
172981 gagaatcctc accccaaaa acatccaaaa aaaggaacca aataataaca aggaagttca
173041 taaaaggatc agatccaaaa taaatagaga aactcatata tctggaaacg caaacatggg
173101 aacatccaa gcgagagatt tagttttttt ttgtttttttt ttttggagaca ggtctcg
173161 ctgtcgccca ggctggatgt cagttgtgc agcgatctca gtcactgc ggctccgc
173221 cccgggttca cgcatttcc ctgcctcagc ctcccgatg gtcggacta caggcgcgc
173281 ccaccacgccc cggctaattt ttgttat tttttagatgac ggggttccac cgtgttagcc
173341 aggtggtct cgtttccat acctcgtat ccacccgcct cggcctccca aagtgttgg
173401 attgcaggcg tgagccaccg caccggcaaa gagatttatt tcagaaggac tttttgtgat
173461 tatctgtgtt tcattttttt gccattctatc taacagagag actgttttc taagacatgg
173521 gaagatgcag gtcattggaa gcaaggatc tgaaggcagg gtcgttgcca ggaaggccgc
173581 gccttgagga cagcaatgtt gggcttggg ctggccaggtaaccatggg tatcaagtgg

173641 ttctgtggcag ctctgcacat gcatgttatga tttgactaa tggtcattt ttaactctgt
173701 ttctccctg ggaataaaga tgtctaaatt ctgagaatat cattagtctg aaagaagcag
173761 gtcggcaggta ttctgagtaa ataatcacct actgtattat aaattaaaag tgcccaga
173821 taacaatctg gtgtaaaata taatggaaaa gaccctattg acaatggcaa taacattcta
173881 agtctcttag taataaacct actggctat gttttaaac atctgaatca atcttcttat
173941 ttccgtgaga aaattaaaca tcatgaagat gtcgttattt tgccagcaa agatttgctg
174001 agggccttct gcaaaggggc cactactta ggcattttt taaaacctca cgatgggtat
174061 acctggccct tcagatgtca gaacatattt tcaaccaaca gcaattaaag accccctggac
174121 tggtgcaaaa gaaaaacaga aaaggccagg cacagtggct catgcctgca atcccagcac
174181 ttggggaggc caaggaggc ggatcacgag gtcaggagat ggagaccatc ctggctaaca
174241 cggtaaacc cggctcttac taaaataca aaaaatttgc cgggtgtgg ggcgggccc
174301 ttagtcccc gctacttgc aggctgaagc aggagaatcg cttgaaccca ggaggcagag
174361 ctgcagtga gccgagatg tgccactgc ctccagctg ggcaacagag tgagactccg
174421 tctccaaaaaa aaaaaaaaaa acaaaacaaa acccagaaaag gcccaaaaata aaccccatat
174481 cttccatgaa cttcacatgat gatggaatat acattccaa gtcagaaga aagagtgaat
174541 tagcaaaccg atgggttga gacaacagca tggcaattt gaaaaaaatg aacacttaaa
174601 accaacaccc tatacagtat aagaaacaca tttcaacttga attaaaggag ttacatattt
174661 aaaaatcaag tcaggaaaaa agcaaaacat ataatggagt atttcttga aaggctgata
174721 ctttctaca ggctgagtaa ccagaagaaa gcaacttagag aggctgacag agctgaccga
174781 agaaacaaaaa caccaggccg gcagaataaa aaaggaagct ctgaaatata aaccccaccc
174841 aagcttagcag agggtaatg actagcccc gggagaacca atcagttca gaggaggaaa
174901 gtggggggc gtaaccacca gccattaaag aaacaagttt caggccggc gtggggctc
174961 acgcctgtaa tcccagact ttggggcc gaggccgggt gatcatgagg tcaggagatc
175021 gagaccatcc tggctaaacaa ggtggaaacc cgtctctact aaaaatacaa aaaattagcc
175081 gggcgcgggt gccccccct gtagtcccc cttactcggg ggtggggc ggagaatggc
175141 gtaaccccg gaagcggagc ttgcagttag cccagatgc gccactgcag tccgcagtcc
175201 ggcctggcg acagaggcag actccgtctc aaaaaaaaaa aaaaaaaaaa agaaacaag
175261 tttcaacaat aacacgattt cctgactccc accaagggtc ccaggccagc ccagccccag
175321 cggactctg caggccttgg ggagccaggt gggcacctgc aggagggtt gggggaggt
175381 gtgaggactc tgagccccca gagtctatca ggacaggctg acaccaagac agagacggcc
175441 tacatctacc caccctatac catgcccagt cctggaccag ccattccctgg ccccaaaaggc
175501 cacagcttag tcaccacagc tgagactgcc ccaatcgagg ccacctgca tgcagcag
175561 cttccctgtca actggcccg ctttgggttt taatgtccct agtcccggaa atgcagaagc
175621 ttcgaaaga gaagaaacgca gcaactctgg caacctgtact gtggggcag acacctggca
175681 cacatcccc cttccccagt gcccgtgatc tccacccagc gctgggcacc ctccaaacctc
175741 cactccattt actccctagac ggagcaaaagg ctggcctcat tggccaccc tccccagcaca
175801 ccacctgac aacccccca gcaagccctgc tcacccggcc accctccccca gtgaggacac
175861 agccctgtac agcctggca ccccaacagg aaagacagcc caaaggccaa gccgcctat
175921 ctgagggttga ggtatgtctg gcaggcagg gactcagct tgaccttgt tgctagaaaa
175981 tgcaccaaga actcagtagc aggccaagga gaggagacat gaattcccc tttcagcaat
176041 tcaccctgtc tgtacaggtg gggagttatc tgtacctctg aggaggtaa gatttcatt
176101 ggtgctccag ctgcagttaga aggagctggc tctgggtct gttggaccc ggttgggtca
176161 tccccatcgc tggctccca tttcccccgc gacattttt atgctcggtt gttggttccc
176221 agaatgtacc aggaaggtagc aggtgcagag tgcagctgac gacgaaggcc ccccaagccag
176281 gaagggtccc caggactct ggctgatgaa agactctggg aaagaaagggt gggaaacagg
176341 agagacaga agggcagggg agggggagagg accataaaggg agaatggaga gaaaaggggaa
176401 aaaggggaga tgagggagaa gagaagggg gggagtgaaa aaaagaggaa tgagaaaaa
176461 agggggaga gagaggagg aaggggagga gaagctggag ggagaggagg gggaaagggg
176521 ggagaggagg gggaggagg ggagaggagg gggaggagg ggagaggagg gggaaagggg
176581 ggagaggagg gggaggagg ggagaggagg gggaggagg ggagaggagg gggaaagggg
176641 ggagaggagg gggaggagg ggagaggagg gggaggagg ggagaggagg gggaaagggg
176701 ggagaggagg gggaggagg ggagaggagg gggaggagg ggaggggctc ccagcaccag
176761 gacccagcac aaccgttca tgacagttt ctctgcagct gcatttaggg ctgttctgga
176821 aactcaggcc ttctccctgt cggcaggctt gggccctgtt ctgggaccc cccttcttct
176881 atcagagcccc ccaggctggg tgaggccac tggatctgt tttactcac aatgcttctg
176941 acaccaaaaca tgggtttcc ctcacacaa ccaggcttcc aacccatccaga cgttagctgg
177001 gtggcccaaca gttgagtgca atgctcacgc caactaccc gggcagctt agccctcagg
177061 ttaaggcttc agtccccacac agcagcccc caactaccc gggcagctt caagttccag ctttccaccc
177121 gtgcctctga ccaggccgt ataaattggg cgggtcccaac aactccctcc tcaggttctg
177181 taatttgcta tttaggttggt gcaaaagttaa ctgtgaccc ggaatcagac agacatgtgt
177241 tcattaatag aacagatcac agaactcagg gaggcattt acttactatt tccagttgtat
177301 tattaaagac acaaccagg aacagctgca tggggggat gggcaggggca aggttaacggg
177361 aagggtctgca gggcctccac gcccctctg ggcactcacc ctcccaccc tccaccc
177421 tcaccaaccc aaagcttcc aaacccctgtc atttagggg tttcatggaa gttccatgac

177481 tccccatcccc agaggtggg gaagtgggc tgaattctca acccaacaat ccagaaaactc
177541 ttatatcagg aacaaggggg ctcagaccaa atctgataat aaaaggtttt tctatcaccc
177601 ttaacactca gggaaattaca aaggtttag gagctctgag gcaggaacca gagatgaaga
177661 acaaacatac atttataatt tcccactatc acagggttcc agacagtacc tcagtctgt
177721 acccaacactg cattgcccc gaagccaga cctcagctca gggccagtgt gctgggatca
177781 gtggccagtg gccagccatc gtgcagagag gggagggcag aggggaggga gggaggggat
177841 gggagaggaa gggagggagg agtgcagaga ggaaggggagg ggaagggagg gaaggagagg
177901 agaccaggag cctcccaggg cctgggacc tgccctggag ggcactgagg agcacaagg
177961 aggatgccta ccactgcccc caccttgcgg tgatcagggtc tgcaaggatca tgctcaacg
178021 cccatggagc tccaggtggc cgccaagtgt ggacagccct gcagcaaagg ccagctactg
178081 cccctggagc cctcacccca ttatctgcg cgcttgcctg tgaggaccc tccctactgt
178141 gtggccttg ggatagccat ggcacggcg ctgccttct tctgggggg ctctgaattt
178201 ccagggtgaga gacccggagag cagaggccca ggagccgaca ccgtgtggc agcttccgc
178261 atcccaagcc gggccatgcc cactggccat cgcttcctt caccctcaagc ccccatgga
178321 atcccttgc actgaggcca gccagcctca tccccctctg ctccctact gacacacacc
178381 ttgtgaggca ggtgggttct taactggaa aatggagggt cacaggccag caagagaaga
178441 gccaggcccc aagtctggtc tggagccccc cactccccac agccctcgct agtgagcc
178501 cgtcatgctg tacaaaagac tctctgtgt agctttgaa aagtacagac ccagggtc
178561 acgcagggtct gataaaagcag aatctctcag gatggggctt atgtacagat gctgtgcaaa
178621 gtcctcaaa tcccatgccc atcagtggc aggccagagag atcacacaat gaaatgctc
178681 ttctgcaaga caagccacag atttgggtt tttctgttt ttttttttt ttttttttta
178741 gacagagtct cactctgctg ctggctgga gtgcagtggc gcaatttcag ctcaactgca
178801 gtcggcctc ctgggttcaa gagattctcc tgccctcagcc tcccaagtag ctgggactac
178861 aggctgtgc caccatggcc agctaattttt tgattttttta atggagatgg gttttcacca
178921 tattggccag gtcggctca aacacctgac ctcagaatgt ctggccgcct cagccctc
178981 aagttacaca ggcgtgagcc actgtgcccc gcctacagat ttcttctaag gaggaaatg
179041 tctttgctt gaaaaacacc ttctagaagc acgtgggata acagataaca gtgggactgg
179101 aagagcagga gactctgcag acagtgcac cacctcagag cagcgggcaaa agtccgctgt
179161 taagacacag gaggagaatg gaggattgc tccagagaaa ctggacagag ggagggagac
179221 gcccactgggaa gggaaagaag aacccacaga tggaaatgtt aatggccccc cacatggcc
179281 aaagggtctt ccaagaacag ggctctgttc ttccagaatg caccgcgtg gccgggaggt
179341 gctggctccc ccacccctctg caggaaggctt gcccctccct acggggctt cccatgggtc
179401 gtaaatgggg actgggttcc atccctggcc cagcaatgt gacagctgag gccgaggctg
179461 aaaaacacaaa caccggaa tcgcaacacc tctggctttt ggaggtgaga aaggggagcg
179521 ttaaggtt aagacacattt tatcttgatt tcagtcctt tactgataaa aatgggatct
179581 tagaaaacga gaattcagag gatattgcct tcacatgaaa tccatctta tccacagca
179641 gcgcgttatt aagcagagaa ataccgggg cgttgcgtt aaagccaggc acagatgagg
179701 atgcgtgttag ccacggctgt gacttaacat cgctctgaag gttccagacg gtgcagattg
179761 gagatgagga tccaaaatta tcagtgat tcaatgtt cttttttttt agaaaaccca
179821 agggactcca caaagatata tttagattaa tgagttcaat caagctgtt gataaaagat
179881 agttctataa aaatcgttag ccttcctgtg tccaataga tggaaagtat tcaatgagca
179941 atggcaaaaaa gaaaaaaaaat ccttggaaata atccgttatt taaaaatgg atgacctgt
180001 taaaatatg tatataatc actctatggt acaacatcaa agcagaatga aataagctga
180061 aaaaatatac tatgttctgt atgaaatactg taaaattatgt caatttcctcc ttaattatgt
180121 ttcttaagttt aatttactat caatcaattt caccgggat ttttcttaccc gacaaaacaa
180181 ttataaaattt catttggaaag aataaaacata tggaccagc caagaaaatt ttgtaaagaga
180241 ataccagcaa agggggacgt atccagccag atgttccaaa gcccacacagt ccaacagaaaa
180301 tagaactcga gctttatatg tagccatgtt tttttttttt aaagacacag gtggaaattaa
180361 ttccagaaat atattttttt taacccagta tgcccaaat attaccattt caacctgtaa
180421 tcaatacttt aaaatgatct attaaacattt ccacattttt tttgggggggt acgaagtctc
180481 tggaaatccaa tgtatattttt atactcaggg tacctctcaa ttccagaccag acacatttt
180541 agtcttcaat agccttgggt ggcttatagc tactgttattt gacagcaag cttcaaaagta
180601 ctttaacttctt tgactttagt agccttacag ttcttaataat aaacacagct taatatttgt
180661 acaagaacag aaagatggaa aagttagatg tatctgggtt actgtttaac ctgagggttc
180721 atttgcactt gactggcgat ggtgtactat aaccctaaaa atagacagat gtaaaatacat
180781 accaatatgc tcacagctt caagccctt ttcatcatat gtccagaagc tctgtcccaa
180841 tcagaatggc cacttagcaag ctttagacaa gccttgcattt aaacgttaag acaccctact
180901 ttccatgtgc tcccaatccc tgcttcaaa cagccgcgtg ggagggatgg gcagggcaag
180961 gtaacgggaa gggctgcaga gcctccacgc cctctctggg cactcaccct cccagcacct
181021 cacctgttca ccaacccaaa gctctccaaa ctttgcgtt taggggtttt caaaggaaaa
181081 ataccgggt gtaaagatga caggctttaa cacttggaaa tactttat acattgaaat
181141 ttccatgttca gggcttccaaa ggaaaaatac ccaggtttaa agatgaatcc cgacattct
181201 gcacagagct ggtggggctt ggtcccttcca acacgcacatc ttccatttcc caacccttga
181261 actgaggatg caatcgccgg gtctcaggcc cccggccctc ctggccacg

181321 acacttccaa cgcaagtccct cactgtcaac accctggaca ctcaaaatgt ctttgaagag
181381 tcacgctgtg aggtcaaaaa gtaggggtgc acgcagaaaag gcgcattgtg ctttgcattg
181441 gtgaggcaaa tacctatccc tggaggaagg accgctactt cgtaattcac aagaggaaga
181501 cccccggagg cgacgaaac caaccgaact cagttcccg gagccgtac aagcgacggc
181561 caggcagcag aaatggcca cccattaggc tgcaagacagc atcagagaaa ccaagcactg
181621 tggcagtggg gcccgcct ccaggaaacc acaaggcaca gagttgcagt ttaaatggg
181681 gttgtcttc tttcttggg gctataaaag taatgggtg tcctgttaacc gatggcacct
181741 tagagcta at gatatgcaga cacggccacc gcttctcac cccgggcct ggggttaggc
181801 ctgcgtcacc acccaccctt tcactgccc tgaagtgcgt gcccacagaa cccatcttt
181861 gctattggag gaactaagg tctcaggatc gtgggtctt tccgcctgc cggtcacgag
181921 gggcacagcg gaccatgtga cccgcaacacg cctccagccc acaataggcc cctgccttc
181981 ctgggttaac cctgtaaatg cacaaggcag ggtggggatg gctgttagcc tgcctggctt
182041 cggcggcacc caggccctgg tcctggacac tagggtttcc tgcgtctgc agcggaaagct
182101 gtgcacccca ctctgcagcc ctggagtctg cgtcgttgt aattaaattt catggacaag
182161 cccgaggagc cagctccggg ctggtaata acttctccat attaatggca gcaccgcaga
182221 cccgtggac agctgatatt taatatgaaa catacgcac agagtgtca gctcctaccc
182281 ggatcctccc attccgggtaa aaccaccagt aggggatggg ggtgtctaac gtggcgccaa
182341 gattgagggg gctggggtcc tcctctcccc ttagtaccc aaccccaatt aaggccccctc
182401 ttctacagag tgcagggaaac ccctcccgag aaatacataa acacatgtc gcccatgtc
182461 actcaaattgc acaccctccc atgcagacac actcatgtc acacacgtgt acacactcat
182521 tcagacgtgc acaaatacatat gcacccccc acgcacactg catgcacgtt ctcctacatg
182581 caagacatgc agggcaaca cgtacatgc caacactcccg ccatgcacag gcccaggaca
182641 cgcacacgtc cacatgcgc caccaggctg ctgtggccgc agcaacacag ggaccccccag
182701 ggtgtcccaa gccagaggg cagggtgggtaa aagaaggac atgtgtgc gccagcaca
182761 acccagggtc ccaggagggtc tgccatgc acctggctt ttgggtctt agataccaca
182821 gtggcaggc ccagaccata gtcaggcag ccacgagca aagccagct cgctgtgag
182881 cgatgtccc tggctgattc cttactcac ccaggtgtgg cccctgtqag cttcacagca
182941 ggcctgggtg aggacctggc cttcacctcc cagctgactc tgcacagggg cctgaattca
183001 ctcactcacc caggtgagc ccctgtgagc ctcacagcag gcctgggtga ggacctggcc
183061 ttacactccc agctgactt gcaggaggcc ctgcattcac tcaggcatga ctgcataagc
183121 actggagaca tgggggacag ggaagccacg gcgagccatt gggcaggaga agcctgcccac
183181 acctgaatga gcacatggcc ggcacattt accagtgc aaaaacatacc cagccacaca
183241 atggagtgc gctctcccg ggagagctgg cttgggtggg ctccggctga gaccagtgt
183301 gactatgcgc gccttcata ggggtggg gccggcttca ctctggccct ccctaaaggaa
183361 ttccaggaga gacgccagcc cacatggaga acctgcataat gtaactccta gaggctcggt
183421 ttccctacctg taaaatgggg ccacgggtgc tggctcacag gaaaaccgca gggactgggt
183481 tcaaageccc aagcagaggt agatgggtgt gatgctgtc ccccgacccc ctccgtggc
183541 tggccggcgc ctccagaggg actgcgtgc actgctggat ccctcaggca tctctgttag
183601 gcccctcagc tcacgttaagg cagcagggg gggagtcat gccccaccct cccaggggca
183661 cctgagggga acgctggagt tctgcctcc tcaatccctt ctccggagt ctcatccag
183721 accagcagag caggcgtactc aggccggagc tggaggggcc tgaccaggc acggggccct
183781 tgaatgcag gccagtggtt ttggatgttt ttaagagcag taagaaccac tgagaggct
183841 taatgtaaagg cgtacatgc gttgtctgc gttttctaaa aatccctctg gctgcagagt
183901 ggagaagggt tgcggggag agaccatggg gcaatgggg gtaaaggc agaaaaggcagg
183961 gttggcacc aggcgtggag ggaagccggg ccctgagaga gctctagaat tggagctggc
184021 taggggaggg cgatcctgg aagggtgggtt ctggaggaga gttggggagag ccacgtgact
184081 ttagtatggg gtcattcagg gaagggtttt ccagaggcag ccggcaggc aaggctggag
184141 cccaggagga cttctaagcc caacatgggg tctggagtc aggagaaggg tggccacaac
184201 cgtgttccacc accctgcaca attctctcc ttctttgtat gataggcccc tcttcattcc
184261 agagtttggg agtctgtccc tccccctcag tcccaaggt gggcaggaga cctgggctag
184321 gcccgcacca cagccatggg tttaagatg ggcctccaaat ccctccccca actctcacag
184381 gtgtcacccca gagaggaggg cttccatct caaagggtgc ctccctggaa aacgagttgg
184441 gctggagcc gctggggggc cgggggttga ggggtggaga gcaattctga aatgaacact
184501 gagaggcagg gctgaaaaat ggatagcatg ctctgagcc ctggatttag ccaggcctga
184561 accaacaact cctcaggacc cccagataat tggggccatg ctgcccattt tttgtttcat
184621 ccaatatgag cagggtttc tgcacttta actgagcaag tcttaagggt acaagatagg
184681 cggaggctga gactggccctt agcagctgtg taggctgtt gcaatgggg cccacaggag
184741 cctctgtcc tacctgggg a ccccgacac cctaggaaac ccagaggttc ctgcagcccc
184801 tctttatgtt ctcgtacaac ctctcattgc gatccaggca gtgccaagggt ggggtgggag
184861 gaaggaagac cttaccagcg gcaatgtcaaa atctcagggtt cccctcacc cactgcacag
184921 tcccaggacc cgctcacacc tgcctgtat gctctagcac ctgagccctgc tggagccca
184981 tgggtggcct cgaaggcaggc agcctgtctg ggaaaaatga aactcatctc gatcccttag
185041 gccggcttct gacttcttc tgggtggctt atgacttta atgtccagca agacagagtg
185101 ggtgcctggc ccagacagca gccccgcagg ggagccccag cctgaaaggc agcagccctg

185161 gatttcttct tgccatcaa ccttgggtct ctctgaggct cagccacgc gtggggcagc
185221 cctatctcac agggttgtgg ggttcgcata aagcatccag ggaaggcat acacttcctc
185281 ctttggcccc caggattgga ccaggcctt gagggcgac catctccacc gcccagccca
185341 tgaggccttc atccactgct gcccagcctg gcaggagagg atccccgtgc tgccctgagt
185401 ggggaggcag aggccggcag ggttgcttcc atcccacgga agtcccagct gcagtcccc
185461 accagtgagg acttaggcca cccataggag ctgcagccca gcgaggtca cgagccagtc
185521 aaggccactc agctccagg cctgctcccg gctccactca cactgacgat ggtcaaacc
185581 acaatgagaa cgtctccagg agtcaaataa gaaagggtgg aaatgttga aagaaatgca
185641 tttttaaaat tcttttgta ggcatgcagg aaaaggcaac aaaacatagt ggcatggcc
185701 ttggaatcag atttgtgttc aaatactggc ttcaagctt gaagcaagct atctcgactt
185761 ttggagcctc agtttattca tctgttagaag ggggatgatg tcggatatg ttccgttaggg
185821 tctgtgtggg attaaacaag gcgagggtggc acagggggtc cagtgagag cctggcacac
185881 aatgttataa ggaacacagt caactgtcac ttaccatcat tacctttac gccaaaaggc
185941 ttctactgg gtttttaatg ggcatgtaaa aaccacgtga aatttagcagc cttaccacgc
186001 tgcccaggag atactgaaa actcttccac agcaacgtga ggcacaacac cactctgtgc
186061 agcctgtttt gctcatgcct tgcgaaagg gacatttgc tggataact atgataacc
186121 cctcgctgtg gatattatta tgaataactt ccgtctagtc actcaaccc tcagagccctg
186181 agtgctccac ctgtaaaccag ggatgagatg cttcccccac cagggctctg tgaggacaag
186241 agtgccctg tgtaaagacc caggccctta gcaaaagaagc ctatctgtgt ggcagcagag
186301 ggggatttaa cagtaggaca gacacatacg cagcaacaaa gtcagaaaag gaaagaagag
186361 gccggcgcg gtggctgagg cctgtaatcc cagcactta ggaggccaag atggcagat
186421 cgcttgaggt tgggagggtt agaccgcctt gatcaaagg aaagaaagaa agaaaaaaag
186481 gaaggaaagga aggaagaaa gagaagagaa gagaaggagg acaagcggaa agaaaagagg
186541 ctggcgcgag tggctcatgc ctgtaatccc agcactttag gaggccgagg caggcagatc
186601 acaaggcttag gagattgaga cccatcttgc taacacgtt gggccatgtc tactaaaaa
186661 attttaaaaa aaataaaaaa aaattagctg cacgtgtgg cgggcccctg tagtccctgc
186721 tacttggag gctgaggcag gagaatggc tgaacccggg aggccagagcg cttgcagtga
186781 gctgagatcg cgccactgca ctccagcctg ggtaatagag tgagactcc tctcaaaaaa
186841 aaaaaaaaaa aagaaagaaa acagaaccaa tggtgcgtc ccactgtcca aggccattat
186901 tcagtgacag cattcaatgc cgctcaggca ggtccagct cctcttccaa ccccaatatt
186961 ctcccagct ctgtctgtt catgggacac ttcaagcaggc ctgtctgtcc caggccaaatc
187021 cacacctggt acctctgcac acgtctccctt ctctgccccg gttggccaaacc tgctgtcttt
187081 ccagctggcc actcattttt ccagccccca ttcaaacacgc ccccccaggag cacactctcc
187141 aagtcttgcg cagatgtcc gtcagccacc ctgacgacag gaggccctggg tgggtaccag
187201 acctgtccca ttgctgacca gtacaggggg taaacatgt agaagggtct ggctgaaaaa
187261 tgaatcacca aatgaactgc tgaatgaatg aaacagtgc aagtctccc caaaatgctc
187321 tccatcagga atgcatagtt cacaggtgtt gggccaccaa cagactgca ccccatgtca
187381 actcccacca ctttaccta ccctaagtcc gcagcaaaaca tctctttca ttttagaaat
187441 taacctccag cgtaaaatgg gacacaccc tcagtgatg tggggccag gatgtcctca gagacagtgg
187501 gcatgtcaa gggctcacca ttttacccca gtcctccag ggtatattac agttgtttgc
187561 aatatggcca agttctaact aaatccagcc aaaggcagtg attttccaaa gcaacaagtg
187621 atattttctg tgcctccata gaaagaatga cagtggttc agagccctgg tgaattgcc
187681 ggctctgcag actgttagat gggctcagc tgcccagta cgggcgtgtc cagggccac
187741 gggggctggg agagtggca ctaactcggg gcaggagttc cctgatgtc aatttagagcg
187801 gaaccatgg tggcgctt cagatgtctg gccagcacag ccagctgac caccacacac
187861 ccaatccaaa atagaaaact gcaatggctt ctgctgtgt ccaggccccg ggcagccgga
187921 cagagcaccg ccaatgtcag cccatctccct cagcatcgat ccaggcttcc tctgaggccc
187981 ctttgcgtc agccacggc cttctgtccc tctccagaa gcagagcttc ccagagctcc
188041 agtcacagga cttagagggc gcaggcagtc tcaaaaggga tcccctagcc tctctcttgc
188101 cttcaggcg ctgccccca aagatcgagc agctcatggg cactagagcc ttcatttcca
188161 aatgcaccccc tcagctaaa atagagaaaa ccaacttttca aaagaaccag ctggagttca
188221 acaggaaacc acagcagaaa aaccaatgtat tcccggctcag ccaaggggca gagggtcaac
188281 tggcatttc agcatctgc ccagctccct gggctgttat taccgcacg caaggaaaca
188341 ggaagactga ctgcccggc ataagtccac cacacacata tcacccaggat atgcacaagg
188401 accgacacaa actcaacaga tgcataacac actgtatgtc agaaaccag acacatacgt
188461 gcaacagaga cactggcaca cgcacatgtt cacacagggt caggccacc cagatgcagg
188521 tacacaccaa cgttccacaa cacaatgtca cacactcaca aaggtagaca cacatacaca
188581 gacccacccc cgatgcaag ccagagactt gttcaggcag gcaaaaagaga gaggccacg
188641 cttctgcgtc ctgagcctt tggagacccg ggcgtcccc atgcataaat ggccccctgc
188701 tcagaccatg caaaacaacc tgctttccag ccaagccctt cagttggagc acgaggctca
188761 ctaaaggcagc ggaagggcgc ggaaagtctt cagcgcaggc cttggccctcc caagggactc
188821 ccaggttaacc cttttccaaa acccaggcga ggcagcggg gagaaggag ctagagaagt
188881 gctactcactc cagccatcg caggctgtca ggcacacaga ggaggcttc gctaggccgg
188941 ccacgtctcc gacgttagaga cagctcccgca caggggtc caccgggtg tgcccttct

189001 cgcctgcca ctccatagtg gccccggcg ccacgaggcg ggcgttggc cgcaaggcga
189061 ggtcaggtc tcggccaaag accgtgacat ttagaaagag gtgRctgcca ggctctcct
189121 cgttgcctc gggRaagctc ggggtccgga ccggggccgc cctgcccgt cgtaccctg
189181 ctttggacgt agctgcccac accaYgtggg acaccaagcg gccttggcg tcagtgcga
189241 cgggcaccgc caggatgcgc tccgctccgt gccccagggg cccgcctgca acggaaagg
189301 gcgttagatc ggcggagacc acggagcccc agtgccttag agacccgcgc gcaagccacg
189361 ccccccaga ccccgccccca ctgcgaaggg aaggggcatt ccggcaggcg accccagaag
189421 ccagctgca cttccccggc tttctgcaa

[0254] Following are human cDNA sequences for transcript variant 1 (long form) and transcript variant 2 (short form) of *ADAMTS2* (cDNA sequences 1 and 2, respectively). Alternative splicing of the *ADAMTS2* gene generates two transcript variants, therefore, *ADAMTS2* exists in two forms: a "long" form comprising a molecule approximately 130 kDa in length, and a "short" form comprising a molecule approximately 70 kDa in length.

ADAMTS2 cDNA Sequence 1 (SEQ ID NO. 2)

NM_014244 Homo sapiens a disintegrin-like and metalloprotease (reprolysin type) with thrombospondin type 1 motif, 2 (ADAMTS2), transcript variant 1, mRNA

1 atggatccgc cggcgggagc cgctcgccgc ctgctctgcc ccgcgtgtct gctgctgctg
61 ctgctgtgc cgccgcgcgt cctggccgcg ccgcgcgcgc cccgcaacgc caggctgcgc
121 gccgcgcgcg accccccagg cggggccctg gggcacggag cggagcgcac cctggcggtg
181 cccgtgcga ctgacgcccc gggccgcgtt gtgtcccacg tgggtgcggc agctacgtcc
241 agagcagggg tacgagcccg caggccgcgc cccgcgcgtt ccccgaggg
301 aacgaggagg agcctggcag tcacctctc tacaatgtca cggcttttgg ccgagacctg
361 cacctgcggc tgccggccaa cggccgcctc gtggccccc gggccactat ggagtggcag
421 ggcgagaagg gcaccaccccg cgtggagccc ctgctggga gctgtctcta cgtcgagac
481 gtggccggcc tagccgaagc ctccctctgtc ggcgtcagca actgcgtatgg gctggctgg
541 ctgatccgga tggaggagga ggagttcttc atcgaaccct tggagaaggg gctggccgc
601 caggaggctg agcaaggccg tgtgcattgtg gtgtatcgcc ggccacccac gtccctcct
661 ctcggggggc cacaggccct ggacacaggg gcctccctgg acagcctgga cagcctcagc
721 cgcgcctgg ggcgtctaga ggagcacggc aacagctgca ggcggagggc acgcaggcat
781 gctgcagacg atgactaca catcgaggc ctgtctggcg tggatgactc tgggtgcag
841 ttccacggga aggacacgt acagaagta ctgtgcacat tcatgaacat tgcataatgaa
901 atctaccatg acgatccctt gggcccccac atcaacgtgg tccctggtgcg gatcatcctc
961 ctgagctatg gaaagtccat gaggcctcatc gagatcgggg accccctctca gagcctggag
1021 aatgtctgcc gctgggccta cctccagcag aagccagaca cggggccacga tgaataccac
1081 gatcacggca tcttcctcac acggcaggac ttggccctt cccgcacatgc aggctatgt
1141 cctgtcaccc gcatgtgcca tccggccgc agctgcaccc tgaaccatgaa ggacggcttc
1201 tcctcagctt ttgtggtggc ccatgagact gcccacgtgc tggcatgga gcacgacgg
1261 cagggcaacc gctgtggcgca cgaggtgcgg ctggcagca tcatggcgcc cctgggtcag
1321 gccgccttcc accgcttcca ctggtccccgc tgccgcgcg aggagctgag ccgcctacctg
1381 cactctatg actgcctgtc ggtgacccccc ttccgcgcgc actggccggc gctgccccag
1441 ctccggggac tgcactactc catgaacggag caatggccgc ttgtacttgg cctgggctac
1501 atgatgtgca cggcgttccg gaccttgc ccctgcacg acgtgtggtg cagccatcct
1561 gacaaccctt acttttgc gaccaagaagaa gggcccccgt tggacgggac tatgtgtgca
1621 cctggcaagc attgtttaa aggacactgc atctggctga cacctgacat cctcaaacgg
1681 gacggcagct gggccgtt ggtccgtt ggcttctgtc cacgtacctg tggcacgggc
1741 gtgaagttca ggacccgcca ggtgtacaac ccacacccgg ccaacggggg cccgcacctgc
1801 tcggcccttgc cctacgactt ccagctctgc agccgcagg actgccccga ctccctggct
1861 gacttccgcg aggacactgtc cggccagtg gacctgtact tggagcacgg cgacgcccag
1921 caccactggc tgccccacga gcaccggat gccaaggaga gatgccaccc tttactgcgag
1981 tccagggaga cggggaggt ggtgtccatg aagcgtatgg tgcgtatgtt gacgcgtgc
2041 tcctacaagg acgccttcag cctctgtgtc cggccggact gacggaaagg gggctgtgac
2101 ggtgtatcg gtcctcagca gacggaaagac aagtgtggc tggcgagg ggacaacagc

2161 cactgcaaag tggtcaaggg cacgttcaca cggtcaccca agaagcatgg ttacatcaag
2221 atgttgaga tccctgcagg agccagacac ctgctcattc aggaggtaga cgccaccagc
2281 caccatctgg ccgtcaagaa cctggagaca ggcaagttca tcttaaatga agagaatgac
2341 gtggatgccca gttccaaaac cttcattgcc atgggctgg agtgggagta cagagacgag
2401 gacggccggg agacgctgca gaccatggc cccctccacg gcaccatcac cggtctggc
2461 atcccggtgg gagacacccc ggtctcactg acgtacaaaat acatgatcca tgaggactca
2521 ctgaatgtcg atgacaacaa cgtcctggaa gaggactctg tggcttacgaa gtggccctg
2581 aagaagtgtt ctccgtgctc caagccctgt ggcggagggt cccagttcac caagtatggc
2641 tgccgcccga ggctggacca caagatggta caccgtggct tctgtgccgc cctctcgaag
2701 cccaaagcca tccgcagagc gtgcaaccca caggaatgtc cccagccagt gtgggtcaca
2761 ggcgaatggg agccatgttag ccagacctgt gggcggacag gcatgcaggt gcgtccgtg
2821 cgctgattc agccgctaca cgacaaaccc acccgctccg tgcacgcaca gcactgcaat
2881 gacggccggc ccggagagccg ccggccctgc agccgcgagc tctggccctgg tcgttggcga
2941 gcccggccct ggtccctgtc ctgcataacc tgtaggaaacgc acggccaggaa gcggccagtg
3001 ccctgcccga ccgggacga cagttcggc atctgcccagg aggagcgtcc tgagacagcg
3061 aggacgtca ggcttggccc ctgtccccca aacatctca atccctccaa gaagagctac
3121 gtagttcaagt ggctgtcccg cccggacccc gactcgccca tccggaagat ctgcgtcaag
3181 ggccactgccc aaggcgacaa gtcaatattc tgtaggatgg aagtcttgc ccgctattgc
3241 tccatcccaag gctacaacaa gctgtccctgc aagtctgtc acctgtacaa caacccatcacc
3301 aacgtggagg gcaggataga gccaccgcct gggaaagcaca acgacattga cgtgttcatg
3361 cctaccctcc cagtggccac ttagccatg gagggtggc catccaag caccggctg
3421 gagggtccctc tcaatgcctc cagcaccaat gcccacaggg atcaccaggaa aaccaatgcc
3481 gtagatgaac cctacaaaat ccatggccctg gaagatgaag tccagccacc caacctaatac
3541 cctcgacac cgagccctta taaaagacc agaaacaaa gaatccaaga gctcattgtat
3601 gagatgcgga agaaagagat gctcggaaat ttctaa

ADAMTS2 cDNA Sequence 2 (SEQ ID NO. 3)

NM_021599 Homo sapiens a disintegrin-like and metalloprotease (reprolysin type) with thrombospondin type 1 motif, 2 (ADAMTS2), transcript variant 2, mRNA

1 atggatccgc cggcgggagc cgctcgccgc ctgctctgcc ccgcgcgtgt gctgtgtgt
61 ctgctgtgc cgcgcgcgtc cctggccgcg cgcgcgcgc cgcgcacgc caggctcgcc
121 gcccggccg accccccagg cggggccctgt gggcacggag cggagcgcac cctggcggtg
181 cccgtgcga ctgacgccccca ggccgcgtt gttttccacg tgggttgcggc agctacgtcc
241 agagcagggg tacgagcccg cagggccccc ccggccgtt ccccgagctt ccccgaggc
301 aacgaggagg agcctggcag tcacctcttc tacaatgtca cggcttttgtt ccgcagacctg
361 cacctggcc tgcggcccaa cggccgcctc gtggccccc gggccactat ggagttggcag
421 ggccagaagg gcaccaccccg cgtggagccc ctgctcggtt gctgtctcta cgtcgagac
481 gtggccggcc tagccgaagc ctccctgtt ggcgtcagca actgcgtatgg gctgggtgtt
541 ctgatccgga tggaggagga ggagttcttc atcgaaccct tggagaaggg gctggccgc
601 caggaggctg agcaaggccc tttttttttt gttttttttt gttttttttt gttttttttt
661 ctcggggggc cacaggccc ggacacaggg gctccctgg acagcctggc cagcctcagc
721 cgcgccttgg cgcgccttgg gggacacggc aacacgtcga ggcggggggc acgcaggcat
781 gctcgacgc atgactacaa catcgaggatc ctgctggcg tggatgactc tttttttttt
841 ttccacggga aggacacgtt acagaaggatc ctgtcgacac tttttttttt tttttttttt
901 atctaccatg acgagtcctt ggggtccctt atcaacgtgg tttttttttt gttttttttt
961 ctgagctatg gaaagtccat gaggctcatc gagatcggtt tttttttttt gttttttttt
1021 aatgtctgcc gctggccctt cttccagcag aagccagaca cggggccacga tttttttttt
1081 gatcacggca tcttccttac acggcaggac tttttttttt cccggcatgca aggctatgtct
1141 cctgtcaccc gcatgtggca tccggccgc agctgcaccc tttttttttt gggccgttcc
1201 tcctcagctt ttgttggcc ccatgagact gggccacgtt gttttttttt gttttttttt
1261 caggcggcacc gctgtggcga cgagggtggc ctggggcagca tttttttttt gttttttttt
1321 gccgccttcc accgccttccca ctgggtccccc tttttttttt gggccgttcc
1381 cactcttgc actgcctgtt ggttggccccc tttttttttt gggccgttcc
1441 ctccggggac tgcactactt catgaacggag caatggccgtt tttttttttt gggccgttcc
1501 atgtatgtca cggcgtttccg gaccttttgc ccctgtttttt gggccgttcc
1561 gacaacccctt acttttgcac gaccaagaag gggcccccctt tttttttttt gggccgttcc
1621 cctggcaagt tcagggccggg cgcgggtggctt catgcctgtt atccacgcac tttttttttt
1681 caaggttaggt ggatcgctt a

[0255] Following are human polypeptide sequences for isoform 1 (long form) and isoform 2 (short form) of *ADAMTS2* (amino acid sequences 1 and 2, respectively).

ADAMTS2 Amino Acid Sequence 1 (SEQ ID NO. 4)

NP_055059 a disintegrin and metalloprotease with thrombospondin motifs-2 isoform 1; procollagen I N-proteinase; Procollagen N-endopeptidase [Homo sapiens]

MDPPAGAARRLLCPALLLLLLLPPPLPPPPPPANARLAAAADPPGGPLGHGAERILAV
PVRTDAQGRLVSHVVAATSRAVRARRAAPVRTPSFPGNNEEPEGSHLFYNVTVFGR
DLHLRLRPNARLVAPGATMEWQGEKGTRVEPLLGSCLYVGDVAGLAEASSVALSNC
DGLAGLIRMEEEFFIEPLEKGLAAQEAEQGRVHVYRRPPTSPPLGGPQALDTGASLDS
LDSLRSRALGVLEEHANSRRRARRHAADDYDYNIEVLLGVDDSVVQFHGKEHVQKYLLT
LMNIVNEIYHDESLGAHINVVLVRIILLSYGKSMSLIEIGNPSQSLENVCRWAYLQQKPD
TGHDEYHDHAIFLTRQDFGPGSMQGYAPVTGMCHPVRSCSTLNHEDGFSSAFVVVAHETG
HVLGMEHDGQGNRCGDEVRLGSIMAPLVQAAFHRFHWSRCSQQELSRYLHSYDCLLD
DPFAHDWPALPQLPGLHYSMNEQCRDFGLGYMMCTAFRTFDPCQLWCSPDNPYF
CKTKKGPPLDGTMCAPGKHCFKGHIWLTPDILKRDGSWGAWSPFGSCRTCGTVKF
RTRQCDNPHPANGGRTCSGLAYDFQLCSRQDCPDSLADFREEQCRQWDLYFEHGDAQ
HHWLPHEHRDAKERCHLYCESRETGEVVSMMKRMVHDGTRCSYKDAFSLCVRGDCRK
VGCDGVIGSSKQEDKCGVCGGDNSHCKVVKGTFRSPKKHGYIKMFEIPAGARHLLIQE
VDATSHHLAVKNLETGKFILNEENDVDASSKTFIAMGVEWEYRDEDGRETLQTMGPLH
GTITVLVIPVGDRVSLTYKYMIEDSLNVDDNNVLEEDSVVYEALKWSPCSKPCG
GGSQFTKYGCRRRLDHKMVHRGFCAALSKPKAIRACNPQECSQPVWVTGEWEPCSQ
TCGRTGMQVRSVRCIQPLHDNTTRSVHAKCNDARPESRRACSRELCPGWRAGPWS
QCSVTCGNGTQERPVPCRTADDSFGICQEERPETARTCRLGPCPRNISDPSKKSYVVQW
LSRPDPDSPIRKISSKGHCQGDKSIFCRMEVLSRYCSIPGYNKLSCKSCNLYNNLTNVEG
RIEPPPGKHNDIDVFMPTLPVPTVAMEVRPSPSTPLEVPLNASSTNATEDHPETNAVDEP
YKIHGLEDEVQPPNLIPRRPSPYEKTRNQRIQELIDEMRKKEMLGKF

ADAMTS2 Amino Acid Sequence 2 (SEQ ID NO. 5)

NP_067610 a disintegrin and metalloprotease with thrombospondin motifs-2 isoform 2; procollagen I N-proteinase; Procollagen N-endopeptidase [Homo sapiens]

MDPPAGAARRLLCPALLLLLLLPPPLPPPPPPANARLAAAADPPGGPLGHGAERILAV
PVRTDAQGRLVSHVVAATSRAVRARRAAPVRTPSFPGNNEEPEGSHLFYNVTVFGR
DLHLRLRPNARLVAPGATMEWQGEKGTRVEPLLGSCLYVGDVAGLAEASSVALSNC
DGLAGLIRMEEEFFIEPLEKGLAAQEAEQGRVHVYRRPPTSPPLGGPQALDTGASLDS
LDSLRSRALGVLEEHANSRRRARRHAADDYDYNIEVLLGVDDSVVQFHGKEHVQKYLLT
LMNIVNEIYHDESLGAHINVVLVRIILLSYGKSMSLIEIGNPSQSLENVCRWAYLQQKPD
TGHDEYHDHAIFLTRQDFGPGSMQGYAPVTGMCHPVRSCSTLNHEDGFSSAFVVVAHETG
HVLGMEHDGQGNRCGDEVRLGSIMAPLVQAAFHRFHWSRCSQQELSRYLHSYDCLLD

DPFAHDWPALPQLPGLHYSMNEQCRDFGLGYMMCTAFRTFDPCQLWCSHPDNPYF
CKTKKGPPLDGTMCAPGKFRPGAVAHACYPSTLGGQGRWIA

[0256] Modifications may be made to the foregoing without departing from the basic aspects of the invention. Although the invention has been described in substantial detail with reference to one or more specific embodiments, those of skill in the art will recognize that changes may be made to the embodiments specifically disclosed in this application, yet these modifications and improvements are within the scope and spirit of the invention, as set forth in the aspects which follow. All publications or patent documents cited in this specification are incorporated herein by reference as if each such publication or document was specifically and individually indicated to be incorporated herein by reference.

[0257] Citation of the above publications or documents is not intended as an admission that any of the foregoing is pertinent prior art, nor does it constitute any admission as to the contents or date of these publications or documents. U.S. patents and other publications referenced herein are hereby incorporated by reference.

What is claimed is:

1. A method for identifying a subject at risk of osteoarthritis, which comprises detecting the presence or absence of one or more polymorphic variations associated with osteoarthritis in a nucleic acid sample from a subject, wherein the one or more polymorphic variations are detected in a nucleotide sequence selected from the group consisting of:
 - (a) a nucleotide sequence in SEQ ID NO: 1-3;
 - (b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-3;
 - (c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-3;
 - (d) a fragment of a nucleotide sequence of (a), (b), or (c);whereby the presence of the polymorphic variation is indicative of the subject being at risk of osteoarthritis.
2. The method of claim 1, which further comprises obtaining the nucleic acid sample from the subject.
3. The method of claim 1, wherein the one or more polymorphic variations are detected within a region spanning chromosome positions 178746000 to 178751000 in human genomic DNA.
4. The method of claim 1, wherein the one or more polymorphic variations are detected at one or more positions selected from the group consisting of 210, 3608, 3609, 4318, 5593, 5629, 5639, 5640, 8943, 17968, 19887, 21034, 21085, , 21596, 23379, 23432, 24007, 26121, 26273, 26755, 27411, 27710, 27842, 28379, 29603, 31232, 31504, 32583, 32794, 32840, 33044, 33150, 33218, 33513, 33959, 34486, 36289, 36570, 38247, 38477, 38518, 38529, 38667, 39781, 39856, 39927, 40506, 41869, 42452, 44788, 46059, 46846, 47712, 48796, 49441, 49602, 49723, 50050, 50171, 50477, 50818, 50833, 50881, 50882, 51386, 51534, 52317, 52368, 52970, 53023, 53356, 53882, 54553, 55475, 55530, 55691, 55848, 55879, 56316, 56911, 57320, 57391, 57437, 57478, 57500, 59111, 59333, 59715, 59804, 59851, 59929, 60052, 60240, 60359, 60381, 60456, 60724, 60875, 60968, 60978, 60998, 61557, 62091, 62645, 62943, 63131, 63145, 63406, 63427, 63554, 63661, 64093, 64153, 64409, 64544, 65257, 65626, 65739 , 66392, 66720, 69177, 69336, 69636, 69823, 69928, 70547, 70633, 71805, 72181, 72200, 72474, 72567, 72973, 73468, 73889, 75730, 75970, 76114, 76342, 76449, 76465, 76791, 78042, 80758, 80778, 81356, 81576, 81689, 81759, 81950, 82562, 83591, 83700, 83821, 83842, 83923, 83929, 84021, 84175, 84417, 84747, 85746, 86129, 86335, 87315, 87648, 87764, 87770, 88221, 90474, 91148, 91150, 91160, 91733, 91772, 91785,

93140, 93148, 96080, 96157, 96313, 96759, 97026, 97320, 97732, 98713, 99707, 99959, 100009, 100020, 100065, 100086, 101270, 101276, 101371, 101376, 101439, 101820, 102392, 102602, 102604, 102896, 189104, 189134 and 189205.

5. The method of claim 1, wherein the one or more polymorphic variations are detected at one or more positions in SEQ ID NO: 1 selected from the group consisting of 5640, 33150, 38247, 38529, 46846, 49723, 50050, 63427, 73889, 189104 and rs428901.

6. The method of claim 1, wherein the one or more polymorphic variations are detected at one or more positions in linkage disequilibrium with one or more positions in claim 3, 4 or 5.

7. The method of claim 1, wherein detecting the presence or absence of the one or more polymorphic variations comprises:

hybridizing an oligonucleotide to the nucleic acid sample, wherein the oligonucleotide is complementary to a nucleotide sequence in the nucleic acid and hybridizes to a region adjacent to the polymorphic variation;

extending the oligonucleotide in the presence of one or more nucleotides, yielding extension products; and

detecting the presence or absence of a polymorphic variation in the extension products.

8. The method of claim 1, wherein the subject is a human.

9. The method of claim 8, wherein the subject is a human female.

10. The method of claim 8, wherein the subject is a human male.

11. A method for identifying a polymorphic variation associated with osteoarthritis proximal to an incident polymorphic variation associated with osteoarthritis, which comprises:

identifying a polymorphic variation proximal to the incident polymorphic variation associated with osteoarthritis, wherein the polymorphic variation is detected in a nucleotide sequence selected from the group consisting of:

- (a) a nucleotide sequence in SEQ ID NO: 1-3;
- (b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-3;

(c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-3;

(d) a fragment of a nucleotide sequence of (a), (b), or (c) comprising a polymorphic variation;

determining the presence or absence of an association of the proximal polymorphic variant with osteoarthritis.

12. The method of claim 11, wherein the incident polymorphic variation is at one or more positions in claim 3, 4 or 5.

13. The method of claim 11, wherein the proximal polymorphic variation is within a region between about 5 kb 5' of the incident polymorphic variation and about 5 kb 3' of the incident polymorphic variation.

14. The method of claim 11, which further comprises determining whether the proximal polymorphic variation is in linkage disequilibrium with the incident polymorphic variation.

15. The method of claim 11, which further comprises identifying a second polymorphic variation proximal to the identified proximal polymorphic variation associated with osteoarthritis and determining if the second proximal polymorphic variation is associated with osteoarthritis.

16. The method of claim 15, wherein the second proximal polymorphic variant is within a region between about 5 kb 5' of the incident polymorphic variation and about 5 kb 3' of the proximal polymorphic variation associated with osteoarthritis.

17. An isolated nucleic acid comprising a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence in SEQ ID NO: 1-3;

(b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-3;

(c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-3;

(d) a fragment of a nucleotide sequence of (a), (b), or (c) comprising a polymorphic variation; and

(e) a nucleotide sequence complementary to the nucleotide sequences of (a), (b), (c), or (d);

wherein the nucleotide sequence comprises a polymorphic variation associated with osteoarthritis selected from the group consisting of a cytosine at position 5640, a cytosine at position 33150, an adenine at position 38247, a thymine at position 38529, an adenine at position 46846, a cytosine at position 49723, a cytosine at position 50050, a cytosine at position 63427, a guanine at position 73889, a thymine at position 189104, and an adenine at position rs428901.

18. An oligonucleotide comprising a nucleotide sequence complementary to a portion of the nucleotide sequence of (a), (b), (c), or (d) in claim 17, wherein the 3' end of the oligonucleotide is adjacent to a polymorphic variation associated with osteoarthritis.

19. A microarray comprising an isolated nucleic acid of claim 17 linked to a solid support.

20. An isolated polypeptide encoded by the isolated nucleic acid sequence of claim 17.

21. A method for identifying a candidate therapeutic for treating osteoarthritis, which comprises:

(a) introducing a test molecule to a system which comprises a nucleic acid comprising a nucleotide sequence selected from the group consisting of:

- (i) a nucleotide sequence in SEQ ID NO: 1-3;
- (ii) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-3;
- (iii) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-3;
- (iv) a fragment of a nucleotide sequence of (a), (b), or (c); or

introducing a test molecule to a system which comprises a protein encoded by a nucleotide sequence of (i), (ii), (iii), or (iv); and

(b) determining the presence or absence of an interaction between the test molecule and the nucleic acid or protein,

whereby the presence of an interaction between the test molecule and the nucleic acid or protein identifies the test molecule as a candidate therapeutic for treating osteoarthritis.

22. The method of claim 21, wherein the system is an animal.

23. The method of claim 21, wherein the system is a cell.

24. The method of claim 21, wherein the nucleotide sequence comprises one or more polymorphic variations associated with osteoarthritis.
25. The method of claim 24, wherein the one or more polymorphic variations associated with osteoarthritis are at one or more positions in claim 3, 4 or 5.
26. A method for treating osteoarthritis in a subject, which comprises contacting one or more cells of a subject in need thereof with a nucleic acid, wherein the nucleic acid comprises a nucleotide sequence selected from the group consisting of:
 - (a) a nucleotide sequence in SEQ ID NO: 1-3;
 - (b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-3;
 - (c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-3;
 - (d) a fragment of a nucleotide sequence of (a), (b), or (c); and
 - (e) a nucleotide sequence complementary to the nucleotide sequences of (a), (b), (c), or (d); whereby contacting the one or more cells of the subject with the nucleic acid treats the osteoarthritis in the subject.
27. The method of claim 26, wherein the nucleic acid is RNA or PNA.
28. The method of claim 27, wherein the nucleic acid is duplex RNA.
29. A method for treating osteoarthritis in a subject, which comprises contacting one or more cells of a subject in need thereof with a protein, wherein the protein is encoded by a nucleotide sequence which comprises a polynucleotide sequence selected from the group consisting of:
 - (a) a nucleotide sequence in SEQ ID NO: 1-3;
 - (b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-3;
 - (c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-3;
 - (d) a fragment of a nucleotide sequence of (a), (b), or (c);whereby contacting the one or more cells of the subject with the protein treats the osteoarthritis in the subject.

30. The method of claim 29, wherein the treatment comprises administration of an effective amount of a composition comprising an active *ADAMTS2* polypeptide or fragment thereof, wherein the polypeptide fragment is selected from the group consisting of: 252-1211, 253-1211, 254-1211, 255-1211, 256-1211, 257-1211, 258-1211, 259-1211 or 260-1211 of SEQ ID NO: 4.

31. The method of claim 30, wherein the polypeptide or fragment has biological activity.

32. A method for treating osteoarthritis in a subject, which comprises:

detecting the presence or absence of one or more polymorphic variations associated with osteoarthritis in a nucleic acid sample from a subject, wherein the one or more polymorphic variation are detected in a nucleotide sequence selected from the group consisting of:

- (a) a nucleotide sequence in SEQ ID NO: 1-3;
- (b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-3;
- (c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-3;
- (d) a fragment of a nucleotide sequence of (a), (b), or (c) comprising a polymorphic variation; and

administering an osteoarthritis treatment to a subject in need thereof based upon the presence or absence of the one or more polymorphic variations in the nucleic acid sample.

33. The method of claim 30, wherein the one or more polymorphic variations are detected at one or more positions in claim 3, 4 or 5.

34. The method of claim 30, wherein the treatment is selected from the group consisting of administering a corticosteroid, a corticosteroid, a nonsteroidal anti-inflammatory drug (NSAID), a cyclooxygenase-2 (COX-2) inhibitor, an antibody, a glucocorticoid, hyaluronic acid, chondroitin sulfate, glucosamine or acetaminophen; prescribing a heat/cold regimen or a joint protection regimen; performing joint surgery; prescribing a weight control regimen; and combinations of the foregoing.

35. A method for detecting or preventing osteoarthritis in a subject, which comprises:

detecting the presence or absence of one or more polymorphic variations associated with osteoarthritis in a nucleic acid sample from a subject, wherein the polymorphic variation is detected in a nucleotide sequence selected from the group consisting of:

- (a) a nucleotide sequence in SEQ ID NO: 1-3;

(b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-3;

(c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-3;

(d) a fragment of a nucleotide sequence of (a), (b), or (c) comprising a polymorphic variation; and

administering an osteoarthritis prevention or detection procedure to a subject in need thereof based upon the presence or absence of the one or more polymorphic variations in the nucleic acid sample.

36. The method of claim 35, wherein the one or more polymorphic variations are detected at one or more positions in claim 3, 4 or 5.

37. The method of claim 35, wherein the osteoarthritis prevention is selected from the group consisting of administering a corticosteroid, a nonsteroidal anti-inflammatory drug (NSAID), a cyclooxygenase-2 (COX-2) inhibitor, an antibody, a glucocorticoid, hyaluronic acid, chondroitin sulfate, glucosamine or acetaminophen; prescribing a heat/cold regimen or a joint protection regimen; performing joint surgery; prescribing a weight control regimen; and combinations of the foregoing.

38. A method of targeting information for preventing or treating osteoarthritis to a subject in need thereof, which comprises:

detecting the presence or absence of one or more polymorphic variations associated with osteoarthritis in a nucleic acid sample from a subject, wherein the polymorphic variation is detected in a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence in SEQ ID NO: 1-3;

(b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in SEQ ID NO: 1-3;

(c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in SEQ ID NO: 1-3;

(d) a fragment of a nucleotide sequence of (a), (b), or (c) comprising a polymorphic variation; and

directing information for preventing or treating osteoarthritis to a subject in need thereof based upon the presence or absence of the one or more polymorphic variations in the nucleic acid sample.

39. The method of claim 38, wherein the one or more polymorphic variations are detected at one or more positions in claim 3, 4 or 5.

40. A composition comprising a cell from a subject having osteoarthritis or at risk of osteoarthritis and an antibody that specifically binds to a protein, polypeptide or peptide encoded by a nucleotide sequence identical to or 90% or more identical to a nucleotide sequence in SEQ ID NO: 1-3.

41. A composition comprising a cell from a subject having osteoarthritis or at risk of osteoarthritis and a RNA, DNA, PNA or ribozyme molecule comprising a nucleotide sequence identical to or 90% or more identical to a portion of a nucleotide sequence in SEQ ID NO: 1-3.

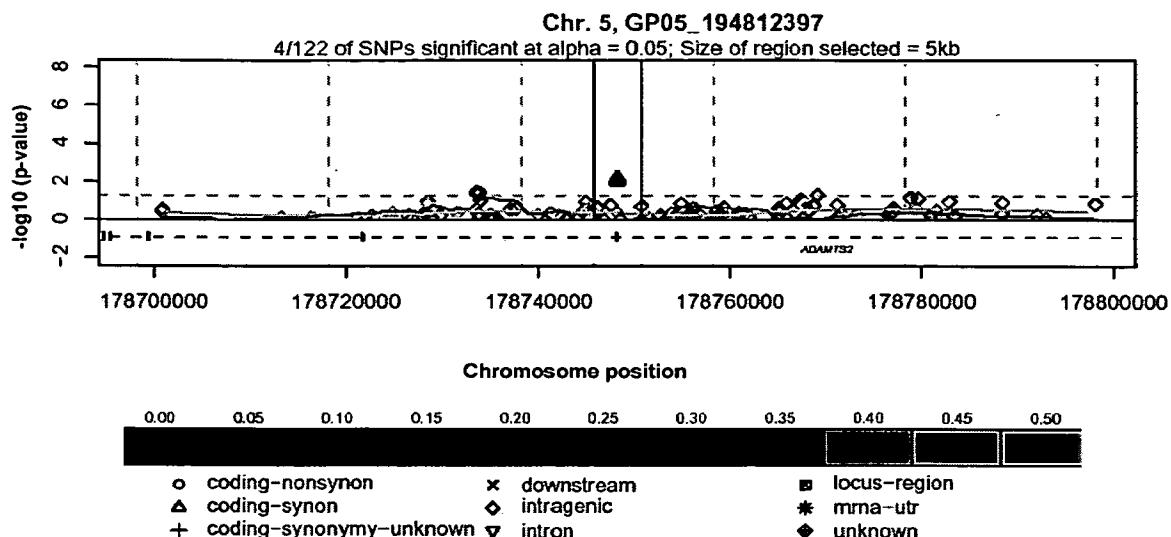
42. The composition of claim 41, wherein the RNA molecule is a short inhibitory RNA molecule.

Abstract of the Disclosure

Provided herein are methods for identifying a risk of osteoarthritis in a subject, reagents and kits for carrying out the methods, methods for identifying candidate therapeutics for treating osteoarthritis, and therapeutic and preventative methods applicable to osteoarthritis. These embodiments are based upon an analysis of polymorphic variations in nucleotide sequences within the human genome.

FIGURE 1

ADAMTS2 – DISCOVERY P-VALUES (female only)



Application Data Sheet

Application Information

Application Type:: Provisional
Subject Matter:: Utility
Suggested Group Art Unit:: Not Yet Assigned
CD-ROM or CD-R?:: None
Sequence submission?:: None
Computer Readable Form (CRF)?:: No
Title:: METHODS FOR IDENTIFYING RISK OF
OSTEOARTHRITIS AND TREATMENTS
THEREOF
Attorney Docket Number:: 524593008900
Request for Early Publication?:: No
Request for Non-Publication?:: No
Total Drawing Sheets?:: 1
Small Entity?:: Yes
Petition included?:: No
Secrecy Order in Parent Appl.?:: No

Applicant Information

Applicant Authority Type:: Inventor
Primary Citizenship Country:: US
Status:: Full Capacity
Given Name:: Steven
Family Name:: MAH
City of Residence:: San Diego
State or Province of Residence:: CA
Country of Residence:: US
Street of mailing address:: 12820 Via Nieve #74
City of mailing address:: San Diego
State or Province of mailing address:: CA

Postal or Zip Code of mailing address:: 92130

Applicant Authority Type:: Inventor
Primary Citizenship Country:: Germany
Status:: Full Capacity
Given Name:: Andreas
Family Name:: BRAUN
City of Residence:: San Diego
State or Province of Residence:: CA
Country of Residence:: US
Street of mailing address:: 3935 Lago Di Grata Circle
City of mailing address:: San Diego
State or Province of mailing address:: CA
Postal or Zip Code of mailing address:: 92130

Applicant Authority Type:: Inventor
Primary Citizenship Country:: Germany
Status:: Full Capacity
Given Name:: Stefan
Middle Name:: M.
Family Name:: KAMMERER
City of Residence:: San Diego
State or Province of Residence:: CA
Country of Residence:: US
Street of mailing address:: 3825 Elijah Court, Unit 334
City of mailing address:: San Diego
State or Province of mailing address:: CA
Postal or Zip Code of mailing address:: 92130

Applicant Authority Type:: Inventor
Primary Citizenship Country:: US
Status:: Full Capacity

Given Name:: Matthew
Middle Name:: Roberts
Family Name:: NELSON
City of Residence:: San Marcos
State or Province of Residence:: CA
Country of Residence:: US
Street of mailing address:: 1250 Calle Prospero
City of mailing address:: San Marcos
State or Province of mailing address:: CA
Postal or Zip Code of mailing address:: 92069

Applicant Authority Type:: Inventor
Primary Citizenship Country:: Sweden
Status:: Full Capacity
Given Name:: Rikard
Middle Name:: Henry
Family Name:: RENELAND
City of Residence:: San Diego
State or Province of Residence:: CA
Country of Residence:: US
Street of mailing address:: 7555 Charmant Drive, #1114
City of mailing address:: San Diego
State or Province of mailing address:: CA
Postal or Zip Code of mailing address:: 92122

Applicant Authority Type:: Inventor
Primary Citizenship Country:: United Kingdom
Status:: Full Capacity
Given Name:: Maria
Middle Name:: L.
Family Name:: LANGDOWN
City of Residence:: San Diego

State or Province of Residence:: CA
Country of Residence:: US
Street of mailing address:: 3701 Yosemite Street
City of mailing address:: San Diego
State or Province of mailing address:: CA
Postal or Zip Code of mailing address:: 92109

Correspondence Information

Correspondence Customer Number:: 25225

Representative Information

Representative Customer Number:: 25225